



The  
University  
Of  
Sheffield.

## Access to Electronic Thesis

Author: Mark Burkitt  
Thesis title: Computational Modelling of Sperm Behaviour in a 3D Virtual Oviduct  
Qualification: PhD

**This electronic thesis is protected by the Copyright, Designs and Patents Act 1988. No reproduction is permitted without consent of the author. It is also protected by the Creative Commons Licence allowing Attributions-Non-commercial-No derivatives.**

This thesis was embargoed until January 2013.

If this electronic thesis has been edited by the author it will be indicated as such on the title page and in the text.

# Computational Modelling of Sperm Behaviour in a 3D Virtual Oviduct

Mark Alexander Burkitt

Submitted in Partial Fulfillment of the  
Requirements for the Degree of  
Doctor of Philosophy

Department of Computer Science

September 2011

Supervisors:

Dr Dawn Walker, Dr Daniela M. Romano and Dr Alireza Fazeli

University of Sheffield



## ABSTRACT

The processes by which individual sperm cells navigate the length and complexity of the female reproductive tract and then reach and fertilise the oocyte are fascinating. Numerous complex processes potentially influence the movement of spermatozoa within the tract, resulting in a regulated supply of spermatozoa to the oocytes at the site of fertilisation. Despite significant differences between species, breeds and individuals, these processes converge to ensure that an optimal number of high quality spermatozoa reach the oocytes, resulting in successful fertilisation without a significant risk of polyspermy.

Computational modelling provides a useful method for combining knowledge about the individual processes in complex systems to help understand the relative significance of each factor. In this thesis, the first agent based computational model of sperm behaviour within the oviductal environment has been created.

First, a generic conceptual model of sperm behaviour within the 3D oviduct is presented. Sperm are modelled as individual cells with a set of behavioural rules defining how they interact with their local environment and regulate their internal state.

Secondly, a set of 3D models of the mammalian oviduct were constructed. Histology images of the mouse oviduct were obtained and the path that the oviductal tube follows through the tissue was identified using CUDA based image analysis. This was used to determine cross-sectional topology, and measurements from the cross sections were used to generate a set of accurately scaled 3D models of the oviduct.

The process of constructing and validating the agent-based computational model of sperm movement and transport within the oviductal environment is described. The model is grounded in reality, with accurate space and time scales used throughout, and parameters and mechanisms from literature where available. Sensitivity analysis is performed on all parameters, and those which are most sensitive to variation have been identified. The model has been validated against literature where possible, and the limitations of the model, validation and assumptions made are clearly stated.

The model has been used to investigate the significance of the oviductal environment on the regulation of sperm distribution and progression to the site of fertilisation, and how changes to that environment alter the distribution. Finally, the potential use for the model and how more complex mechanisms could be integrated in the future are discussed.

## ACKNOWLEDGEMENTS

*First I would like to thank my wife Gaby and my daughter Sophie for their continuous support during my PhD and helping to motivate me when needed. I would like to thank my supervisor Dr Dawn Walker for her guidance throughout, ensuring that I keep on track and focused on the task in hand. I would also like to thank Dr Alireza Fazeli for his enthusiasm and encouragement to actively participate both in his group and the wider reproductive biology community. I would like to thank Dr Daniela Romano for providing encouragement and support during my PhD. I would also like to thank my colleagues and friends, in particular Mariam Kiran, Carmen Almiñana-Brines, Mesude Bicak and Afsaneh Maleki for providing support, advice and discussions about different aspects of my project.*

## DECLARATION

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text. This work has not been submitted for any other degree or professional qualification except as specified.

Mark Burkitt

# Contents

<b>List of Tables</b>	<b>xiii</b>
<b>List of Figures</b>	<b>xv</b>
<b>Glossary</b>	<b>xxv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Overview of mammalian reproduction . . . . .	2
1.1.1 Research objectives . . . . .	2
1.2 Thesis structure . . . . .	4
1.3 Contribution to knowledge . . . . .	5
1.4 Publications . . . . .	6
1.4.1 Peer reviewed papers . . . . .	6
1.4.2 Abstracts, presentations & posters . . . . .	7
<b>2 Reproductive biology</b>	<b>8</b>
2.1 Components of fertilisation . . . . .	8
2.1.1 Spermatozoa . . . . .	8
2.1.2 Oocytes . . . . .	10
2.1.3 The oviduct . . . . .	10
2.2 The reproductive process . . . . .	12

<i>CONTENTS</i>	v
2.2.1 Ovulation . . . . .	12
2.2.2 Sperm progression . . . . .	12
2.2.3 Complex guidance mechanisms . . . . .	14
2.2.3.1 Fluid dynamics . . . . .	15
2.2.3.2 Local maternal communication . . . . .	16
2.2.3.3 Thermotaxis & chemotaxis . . . . .	16
2.3 Species specific variations . . . . .	17
2.4 Conclusions . . . . .	18
<b>3 Computational modelling of biological systems</b>	<b>20</b>
3.1 The complexity of biological systems . . . . .	20
3.2 Computational biology . . . . .	20
3.2.1 Computational systems biology . . . . .	21
3.3 Computational modelling of biological systems . . . . .	22
3.3.1 Modelling complex systems . . . . .	23
3.3.2 Modelling approaches . . . . .	23
3.3.3 Modelling techniques . . . . .	25
3.3.3.1 Discrete modelling . . . . .	25
3.3.3.2 Continuous modelling . . . . .	26
3.4 Computational modelling applied to biological systems . . . . .	27
3.4.1 Agent based modelling of cell behaviour . . . . .	28
3.4.2 Modelling in reproductive biology and related systems . . . . .	30
3.5 Conclusions . . . . .	32
<b>4 Defining a model of sperm behaviour</b>	<b>34</b>
4.1 Modelling process . . . . .	34

4.2	Domain Model . . . . .	35
4.2.1	Defining the system boundaries . . . . .	35
4.2.1.1	Physical environment . . . . .	36
4.2.1.2	Individuals within the environment . . . . .	36
4.2.1.3	Temporal environment . . . . .	36
4.2.2	Defining important behaviour . . . . .	37
4.2.2.1	Capacitation . . . . .	37
4.2.2.2	Sperm movement . . . . .	37
4.2.2.3	Sperm interactions with the oviduct . . . . .	38
4.2.2.4	Sperm interactions with oocytes . . . . .	38
4.3	Defining an agent model of sperm behaviour . . . . .	39
4.3.1	Creating the system boundaries . . . . .	40
4.3.1.1	Physical environment . . . . .	43
4.3.1.2	Individuals within the environment . . . . .	43
4.3.1.3	Temporal environment . . . . .	44
4.3.1.4	Timing of interactions . . . . .	44
4.3.2	Agent behaviour . . . . .	44
4.3.2.1	Capacitation . . . . .	45
4.3.2.2	Capacitated sperm death . . . . .	45
4.3.2.3	Sperm movement . . . . .	46
4.3.2.4	Sperm interactions with the oviduct . . . . .	46
4.3.2.5	Sperm interactions with oocytes . . . . .	49
4.3.2.6	Sperm to sperm interactions . . . . .	49
4.3.3	Inputs and outputs . . . . .	49
4.4	Populating the model with parameters . . . . .	50

4.4.1	Initial conditions . . . . .	51
4.4.1.1	Initial simulation time and sperm distribution . . . . .	51
4.4.1.2	Ovulation time . . . . .	54
4.4.1.3	Number of sperm . . . . .	54
4.4.1.4	Number of oocytes . . . . .	54
4.4.1.5	3D environment . . . . .	54
4.4.2	System level parameters . . . . .	55
4.4.2.1	Sperm radius . . . . .	55
4.4.2.2	Oocyte radius . . . . .	55
4.4.2.3	Capacitation threshold . . . . .	55
4.4.2.4	Attachment and detachment thresholds . . . . .	55
4.4.2.5	Time in progressive and non-progressive state . . . . .	56
4.4.2.6	Capacitated sperm life . . . . .	56
4.4.2.7	Oocyte maturation time . . . . .	56
4.4.2.8	Progressive and non-progressive velocity . . . . .	56
4.4.2.9	Number of movement steps . . . . .	58
4.4.2.10	Rotation and reflection angles . . . . .	58
4.4.2.11	Block to polyspermy . . . . .	59
4.4.2.12	Total simulation time . . . . .	59
4.5	Conclusions . . . . .	60
<b>5</b>	<b>Creating a 3D virtual oviduct</b>	<b>61</b>
5.1	3D reconstruction of biological organs . . . . .	61
5.2	The structure of the mouse oviduct . . . . .	63
5.2.1	Requirements for model construction . . . . .	64
5.2.2	Visualisation of histology images . . . . .	64

5.3	Model construction . . . . .	65
5.3.1	Particle systems . . . . .	66
5.3.2	High performance computing . . . . .	66
5.4	Extraction of structural information from histology images . . . . .	67
5.4.1	Identification of lumen cross-sections . . . . .	67
5.4.2	Curved path identification . . . . .	71
5.4.3	Internal cross-section generation . . . . .	74
5.5	Binding the cross-sections to the 3D curve . . . . .	78
5.5.1	Force definitions . . . . .	79
5.5.1.1	Base forces . . . . .	79
5.5.1.2	Cross-sectional integrity . . . . .	80
5.5.1.3	Structural integrity . . . . .	80
5.5.1.4	Form and realism . . . . .	81
5.5.2	Force resolution computation . . . . .	83
5.5.3	Implementation . . . . .	83
5.5.3.1	Data structures . . . . .	83
5.5.3.2	Force functions . . . . .	84
5.6	Validation . . . . .	85
5.6.1	Internal cross section validation . . . . .	85
5.6.2	Lumen path identification . . . . .	89
5.6.3	Limitations . . . . .	89
5.7	3D models for agent based simulations . . . . .	92
5.7.1	Oviductal regions . . . . .	93
5.8	Optimisation . . . . .	93
5.8.1	Image processing . . . . .	93



5.8.2	Cross-section resolution . . . . .	95
5.9	Conclusion . . . . .	96
<b>6</b>	<b>Computational model implementation</b>	<b>98</b>
6.1	Selecting an ABM framework . . . . .	98
6.1.1	Framework summary . . . . .	99
6.1.1.1	Swarm . . . . .	99
6.1.1.2	Repast . . . . .	100
6.1.1.3	MASON . . . . .	100
6.1.1.4	FLAME . . . . .	100
6.1.1.5	FLAME GPU . . . . .	101
6.1.2	StupidModel . . . . .	102
6.1.2.1	Version 1: mobile agents . . . . .	102
6.1.2.2	Version 2: agent growth . . . . .	103
6.1.2.3	Version 3: habitat cells and resources . . . . .	103
6.1.2.4	Version 5: parameters and parameter displays . . . . .	103
6.1.2.5	Version 7: stopping the model . . . . .	104
6.1.2.6	Version 8: file output . . . . .	104
6.1.2.7	Version 9: randomisation of bug movement . . . . .	104
6.1.2.8	Version 10: size-ordering of bug movement . . . . .	105
6.1.2.9	Version 11: optimal movement . . . . .	105
6.1.2.10	Version 12: mortality and reproduction . . . . .	105
6.1.2.11	Version 14: random normal initial size . . . . .	106
6.1.2.12	Version 15: habitat data read from a file . . . . .	106
6.1.2.13	Version 16: predators . . . . .	106
6.1.2.14	Discussion . . . . .	106

6.1.3	Basic sperm model . . . . .	107
6.1.3.1	Evaluation criteria . . . . .	107
6.1.3.2	Model description . . . . .	108
6.1.4	Discussion . . . . .	112
6.2	Model implementation . . . . .	113
6.2.1	Representation of 3D oviductal geometry . . . . .	113
6.2.2	Agent parameters . . . . .	113
6.2.3	Agent functions . . . . .	116
6.2.3.1	Conic rotation implementation . . . . .	116
6.2.3.2	Collision detection implementation . . . . .	116
6.3	Model initialisation . . . . .	119
6.4	Outputs . . . . .	120
6.4.1	Data management and analysis . . . . .	121
6.4.2	Visualisation . . . . .	121
6.5	Optimisation . . . . .	124
6.5.1	Collision detection . . . . .	125
6.5.2	Concurrent simulations . . . . .	125
6.6	Conclusions . . . . .	126
<b>7</b>	<b>Calibration and validation</b>	<b>127</b>
7.1	Validation . . . . .	127
7.1.1	Levels of validation . . . . .	128
7.1.2	Sensitivity analysis . . . . .	129
7.1.3	Validation process . . . . .	129
7.1.4	Literature for calibration . . . . .	130
7.1.5	Metrics for calibration and validation . . . . .	131

7.2	Model calibration . . . . .	132
7.2.1	Initial simulation runs . . . . .	132
7.2.2	Uncertain parameter analysis . . . . .	132
7.2.2.1	Capacitation . . . . .	133
7.2.2.2	Progression and distribution . . . . .	135
7.2.3	Fully calibrated model . . . . .	140
7.3	Sensitivity analysis . . . . .	143
7.3.1	Sensitivity to system parameters . . . . .	143
7.3.2	Sensitivity to morphological parameters . . . . .	144
7.4	Results validation against literature . . . . .	150
7.4.1	Comparison with literature . . . . .	151
7.4.2	Validation study 1 . . . . .	153
7.4.2.1	Validation Study 1.1 . . . . .	155
7.4.3	Validation study 2 . . . . .	156
7.5	Conclusions . . . . .	158
<b>8</b>	<b>Experimentation</b>	<b>160</b>
8.1	Experiment 1: modifying the 3D environment . . . . .	160
8.1.1	Natural variations between individuals . . . . .	161
8.1.2	Relative significance of bends and folds . . . . .	161
8.1.2.1	Sperm reservoir . . . . .	166
8.1.2.2	Lower isthmus . . . . .	166
8.1.2.3	Upper isthmus . . . . .	167
8.1.2.4	Ampulla . . . . .	168
8.1.2.5	Pre-ampulla . . . . .	168
8.1.3	Penetrated oocytes . . . . .	168

<i>CONTENTS</i>	xii
8.1.4 Polyspermic oocytes . . . . .	170
8.1.5 Experiment 1: Summary and conclusions . . . . .	171
8.2 Experiment 2: sperm population size . . . . .	171
8.3 Conclusions . . . . .	176
<b>9 Conclusion</b>	<b>178</b>
9.1 Limitations and future work . . . . .	181
<b>Bibliography</b>	<b>183</b>
<b>A Micro-CT experiments</b>	<b>204</b>
<b>B Obtaining histology images from mouse oviducts</b>	<b>206</b>
<b>C Oviduct model details</b>	<b>208</b>
<b>D Model assumptions</b>	<b>213</b>

## List of Tables

4.1	Sperm and oocyte specific system level parameters used in the model. .	42
4.2	Agent level parameters used in the model. . . . .	42
4.3	Inputs to the model. . . . .	43
4.4	Initial values of sperm and oocyte specific system level parameters used in the model which were identified from literature for the mouse. The parameters with estimated or uncertain values are marked with an asterisk. (a) Firman and Simmons (2010), (b) Kim et al. (2004), (c) McGaughey et al. (1968), (d) Demott and Suarez (1992), (e) Bedford (1983), (f) Zamboni (1972), (g) Braden and Austin (1954 <i>a</i> ), (h) Goodson et al. (2011), (i) Wolf (1978) . . . . .	53
4.5	Initial values of inputs used in the model which were identified from literature for the mouse. (a) Braden and Austin (1954 <i>a</i> ), (b) Hodges et al. (2008), (c) Zamboni (1972), (d) Nicol and McLaren (1974), (e) Generated as described in Chapter 5. . . . .	53
5.1	Describes the measurements taken from the histology image cross-sections. . . . .	75
5.2	Hardware configuration used for all simulations. . . . .	93
5.3	Time taken to complete processing of 33 images with dimensions 1481 by 1304 using the different numbers of threads. . . . .	95
5.4	Relative number of iterations and time taken for the particle physics simulation to to reach equilibrium for each implementation. . . . .	96
6.1	Hardware configuration used for all simulations. . . . .	109

6.2	Sperm and oocyte agent parameters. . . . .	114
6.3	Hardware configuration used for all simulations. . . . .	124
7.1	The expected total number of sperm at the site of fertilisation after 9 hours <i>post coitus</i> . The multiple measurements from Braden (1958) are from crosses of different strains of mouse. . . . .	131
7.2	Final values of system level parameters used in the model after model calibration. The parameters with estimated or uncertain values are marked with an asterisk. (a) Firman and Simmons (2010), (b) Kim et al. (2004), (c) McGaughey et al. (1968), (d) Demott and Suarez (1992), (e) Bedford (1983), (f) Zamboni (1972), (g) Braden and Austin (1954 <i>a</i> ), (h) Goodson et al. (2011), (i) Wolf (1978) . . . . .	141
7.3	How the relative distribution of sperm changes as the shrinkage is increased. SR = sperm reservoir, LI = lower isthmus, UI = upper isthmus, A = ampulla, PA = pre-ampulla. . . . .	150
7.4	How the timing of HCG injections reported in McGaughey et al. (1968) relates to the time of ovulation, the [Ovulation Time] input parameter and the [Total Simulation Time]. . . . .	157
8.1	The different sperm population sizes used for the experiment. The factor differences from the average population size identified from literature is also shown. . . . .	172

## List of Figures

2.1	The oviduct in relation to the rest of the female reproductive system. The diagram is based on human anatomy. . . . .	9
2.2	A schematic diagram of the basic components of a spermatozoon. . . .	9
2.3	A schematic diagram of the basic components of an oocyte with surrounding cumulus complex. . . . .	10
2.4	The different phases of the menstrual cycle in humans and the oestrous cycle in mammals. The ‘Fertility Period’ relates to the time frame in which coitus can result in successful conception. Based on descriptions presented in Johnson and Everitt (1999). . . . .	13
4.1	The main regions of the oviduct along with an overview of the progression of sperm and oocytes within the system. . . . .	36
4.2	The processes followed by each sperm agent every iteration. . . . .	40
4.3	The different states individual sperm agents can have and transitions between those states. (a) Sperm can switch between progressive and non-progressive movement. (b) Initially uncapacitated, sperm gradually become capacitated then die once their lifetime has completed. (c) Sperm swimming freely can either attach to the oviduct, attach to an oocyte or die. . . . .	41
4.4	Differences between the rotation and reflection methods used within the model. (a) Non-progressive sperm rotate randomly between 0 and $\theta$ and (b) the resulting potential rotation cone in 3D. (c) Sperm interacting with the oviduct rotate $\alpha$ degrees to be parallel to the surface, then randomly between 0 and $\theta$ and (d) the resulting potential rotation half-cone in 3D. . . . .	47

4.5	The timeline of events at the start of the mouse reproductive cycle. The model represents the period of time from when the UTJ closes 1h <i>post coitus</i> and lasts for eight hours. (a) Suarez and Pacey (2006) , (b) Zamboni (1972), (c) Braden and Austin (1954a), (d) Perkins (1972), (e) Harper (1982), (f) McGaughey et al. (1968) . . . . .	52
5.1	(a) A photograph of a complete mouse oviduct encased in fatty tissue and (b) a view of a histology image of the mouse oviduct showing its internal structure. Several cross-sections labelled A - F are clearly visible.	64
5.2	Stacked histology images with transparency applied to the non-tissue regions of each image. The illusion of volume is obtained by closely stacking 200 planes per image slice. The slice thickness is exaggerated from $5\mu m$ to $20\mu m$ for visual clarity. . . . .	65
5.3	The complete cross-section identification process for a single image and how the different processes are distributed between the CPU and GPU. (a) The image is transferred to the GPU, where the tissue regions and boundaries are identified. The red border represents the 1 pixel region adjacent to the block. (b) The non-tissue regions and the id associated with each pixel are combined, then the remaining ids are iteratively reduced to hold the minimum value of all neighbouring pixels. Once completed, the data is transferred back to the CPU. The image shows a representation of the data with each unique remaining id holding a different colour. (c) The identified regions are combined with the border data and converted into a set of lists, one for each cross-section and passed to the GPU. Within each list, N-Body comparison and 1D reduction is performed to determine the largest distance between two points within a cross-section. This is then transferred back to the CPU and used to calculate an ellipse which encapsulates the corresponding cross-section.	69
5.4	The relationship between (a) the elliptical regions identified from the histology images and (b) the final 3D curve after the path has been identified. . . . .	72



5.5	The path connection rules used to determine the cross-sectional topology. (a) A missing cross-section is identified in between two vertically overlapping cross-sections. (b) Clusters of points which are interconnected are merged together. (c) A single cross-section above or below two merging branches (left) is assumed to be the top or bottom of a bend in the tube, so they are merged together (right). . . . .	75
5.6	The different measurements taken from the histology images. (a) The outer wide (blue) and narrow (red) diameter of the tube. (b) The inner wide (blue) and narrow (red) diameter of the lumen. (c) The length of the internal lumen. (d) Measurements from the wide (blue) and narrow (red) portions of the lumen. . . . .	76
5.7	A series of generated cross-sections and how they relate to different regions of the oviduct. . . . .	77
5.8	Cross-sections bound to the 3D curve and rotated to match the curve gradient. Collisions between adjacent sections are red. . . . .	78
5.9	(a) The Spoke and (b) the Radial Integrity Forces which maintain the radial shape of the cross-section and the connection between adjacent and opposite particles. . . . .	80
5.10	The different Structural Integrity Forces. The Loose Integrity Force (a) maintains a connection between particles from adjacent sections. The Crossover Force (b) prevents intersections between adjacent sections. The Gradient Force (c) prevents the tube from flattening. . . . .	81
5.11	The Form and Realism Forces. The Position Force (a) ensures that the mid-point remains close to its corresponding position along the curve. The Smoothing Force (b) removes sharp angles by applying a force to the middle particle when the angle becomes large. The Collision Force (c) prevents different sections of the tube from intersecting. . . . .	82
5.12	A translucent view of the completed 3D model after all collisions have been resolved. (CI) is the caudal isthmus just above the UTJ. (LI) is the lower isthmus. (UI) is the upper isthmus. (A) is the ampulla region. . . . .	84
5.13	Differences between how functions [F0 Fn] are processed on single-core, multi-core and GPU processors. . . . .	85

5.14	Comparison between three generated cross sections A, B & C and three histology images from the corresponding position in the sperm reservoir of the oviduct. Two images from histology set B slice 54 and one image from histology set C slice 34 are shown for comparison. . . . .	86
5.15	Comparison between three generated cross sections A, B & C and three histology images from the corresponding position in the lower isthmus of the oviduct. One image from histology set A slice 14, one image from histology set B slice 74 and one image from histology set C slice 32 are shown for comparison. . . . .	86
5.16	Comparison between three generated cross sections A, B & C and three histology images from the corresponding position in the upper isthmus of the oviduct. One image from histology set A slice 16, one image from histology set B slice 79 and one image from histology set C slice 14 are shown for comparison. . . . .	87
5.17	Comparison between three generated cross sections A, B & C and three histology images from the corresponding position in the ampulla-isthmic junction of the oviduct. One image from histology set A slice 10 and two images from histology set B slices 67 and 68 are shown for comparison. . . . .	88
5.18	Comparison between three generated cross sections A, B & C and two histology images from the corresponding position in the site of fertilisation in the ampulla of the oviduct. One image from histology set A slice 19 and one image from histology set B slice 74 are shown for comparison. . . . .	89
5.19	Comparison between three generated cross sections A, B & C and three histology images from the corresponding position in the pre-ampulla of the oviduct. One image from histology set A slice 22, one image from histology set B slice 64 and one image from histology set C slice 30 are shown for comparison. . . . .	90
5.20	Curves identified from 3 data sets superimposed on their intensity maps. Although not visually apparent from the images, the curves match well to the original image data. . . . .	91
5.21	The different regions of the oviduct and how they relate to the generated cross-sections. The bold groupings indicate where the individual regions are combined for analysis. . . . .	94

6.1	Relative performance of each framework when run for 100 iterations with (a) 1,000 agents, (b) 10,000 agents and (c) 100,000 agents. SW = Swarm, RE = Repast, MA = MASON, MAP = MASON in Parallel, FGPU = Flame GPU, FWIN = Flame in serial under Windows, FG1 - FG32 = Flame run on the Linux Grid using 1 - 32 cores. Error bars show S.D. .	110
6.2	Relative performance of the fastest four frameworks when run for 100 iterations with (a) 1,000 agents, (b) 10,000 agents and (c) 100,000 agents. FGPU = Flame GPU, FG8 - FG32 = Flame run on the Linux Grid using 8 - 32 cores. Error bars show S.D. . . . .	111
6.3	How the capacitation state, movement state, attachment state and alive state are encoded into a single variable. The alive state is implicitly determined by the presence or absence of the capacitation state. . . .	115
6.4	How the potential slices are identified. The initial slice range (3 slices) is progressively extended as intersections with the corresponding slice planes are identified. The algorithm continues until slice plane (d) is processed where no intersection occurs. The final slice range is 6. . . .	118
6.5	(a) An oviduct cross-section with possible initial positions for sperm agents and (b) an oviduct cross-section at the site of fertilisation with possible initial positions for oocyte agents. . . . .	119
6.6	The 3D oviduct model (OvA1) that was used for the initial simulations. The initial position of sperm agents within the sperm reservoir and the oocytes at the site of fertilisation are shown. The displayed size of the sperm agents has been increased for visual clarity. . . . .	120
6.7	(a) The environment with transparency applied and the agents visible through the structure, (b) the agents displayed without the environment, (c) the environment with transparency applied and the radius enhancement of sperm agents and (d) the radius enhancement of sperm agents without the environment. oocyte agents. . . . .	123
6.8	(a) The environment with transparency applied and the agents with lines showing their path, (b) the agents and lines showing their path displayed without the environment, (c) the environment with transparency applied, the radius enhancement of sperm agents and lines showing their path and (d) the radius enhancement of sperm agents and lines showing their path without the environment. oocyte agents. . . . .	124

6.9	(a) The time taken to complete 300 iterations using different collision detection strategies. The scale is logarithmic due to the exponential difference in time between the naïve and optimised approaches. N = naïve, BS = bounding sphere, SR = slice range, O = optimised. (b) the average time taken to complete a single simulation when multiple simulations are run concurrently. . . . .	125
7.1	The processes followed for calibrating and validating the model of sperm behaviour. . . . .	130
7.2	(a) The final distribution of sperm when run with the initial simulation parameter values and (b) the percentage of sperm which are capacitated over time. Error bars show S.D. SR = sperm reservoir, LI = lower isthmus, UI = upper isthmus, A = ampulla, PA = pre-ampulla. . . . .	132
7.3	Analysis of the [Capacitation Threshold] when varied between 0 and 1 in increments of 0.1. . . . .	134
7.4	(a) A heat map of the [Capacitation Threshold], varying (ct) between 1 and 24 in increments of 1 and (b) how the percentage of capacitated sperm at t=2 hours changes as the [Capacitation Threshold] (ct) is varied between 1 and 24 in increments of 1. . . . .	135
7.5	(a) The range and distribution of individual measurements when each parameter is varied through its full range and (b) the difference in the sperm population in the ampulla after 8 hours which can be obtained by varying individual parameters through their entire range. The error bars show S.D. AT = [Progressive Attachment Threshold], DA = [Max. Detachment Angle], NPA = [Max. Non-Progressive Angle], RA = [Max. Reflection Angle], MS = [No. of Movement Steps]. . . . .	136
7.6	(a) The impact that varying the [Progressive Attachment-Threshold] through its full range has on the number of sperm reaching the ampulla after 8 hours and (b) the impact of small changes over the identified subrange. . . . .	137
7.7	The difference in the sperm population in the ampulla after 8 hours that can be obtained by varying individual parameters through their entire range. The error bars show S.D. . DA = [Max. Detachment Angle], NPA = [Max. Non-Progressive Angle], RA = [Max. Reflection Angle], MS = [No. of Movement Steps] . . . . .	138

- 7.8 How variations of (a) the [Max. Detachment Angle] (DA), (b) the [No. of Movement Steps] (MS), (c) the [Max. Non-Progressive Angle] (NPA), and (d) the [Max. Reflection Angle] (RA) parameters influence the number of sperm in the ampulla after 8 hours. . . . 139
- 7.9 The impact of variations in the [Max. Reflection Angle] on the distribution of sperm within the range 35 - 60 degrees. . . . . 140
- 7.10 Final distribution of sperm within the oviduct after 8 hours using the calibrated parameters. SR = sperm reservoir, LI = lower isthmus, UI = upper isthmus, A = ampulla, PA = pre-ampulla. . . . . 142
- 7.11 Results from sensitivity analysis of all system parameters in relation to the number of sperm in the ampulla after 8 hours, with (a) values reduced by 10% and (b) values increased by 10%. Control = The calibrated average, AT = [Progressive Attachment Threshold], CT = [Capacitation Threshold], DA = [Max. Detachment Angle], RA = [Max. Reflection Angle], NPA = [Max. Non-Progressive Angle], OR = [Oocyte Radius], SR = [Sperm Radius]. . . . . 145
- 7.12 Results from sensitivity analysis of all system parameters in relation to the number of sperm in the ampulla after 8 hours, with (a) values reduced by 10% and (b) values increased by 10%. Control = The calibrated average, NPMin = lower boundary of the [Time in Non-Progressive State], NPMax = upper boundary of the [Time in Non-Progressive State], PMin = lower boundary of the [Time in Progressive State], PMax = upper boundary of the [Time in Progressive State], CSL = [Capacitated Sperm Life], NPV = [Non-Progressive Velocity], PV = [Progressive Velocity]. . . . . 146
- 7.13 Results from sensitivity analysis of all system parameters in relation to the number of penetrated oocytes, with (a) values reduced by 10% and (b) values increased by 10%. Control = The calibrated average, AT = [Progressive Attachment Threshold], CT = [Capacitation Threshold], DA = [Max. Detachment Angle], RA = [Max. Reflection Angle], NPA = [Max. Non-Progressive Angle], OR = [Oocyte Radius], SR = [Sperm Radius]. . . . . 147

7.14	Results from sensitivity analysis of all system parameters in relation to the number of penetrated oocytes, with (a) values reduced by 10% and (b) values increased by 10%. Control = The calibrated average, NPMIn = lower boundary of the [Time in Non-Progressive State], NPMax = upper boundary of the [Time in Non-Progressive State], PMin = lower boundary of the [Time in Progressive State], PMax = upper boundary of the [Time in Progressive State], CSL = [Capacitated Sperm Life], NPV = [Non-Progressive Velocity], PV = [Progressive Velocity]. . . . .	148
7.15	The impact of shrinking the lumen to 90% and 80% or its original size when compared to the Normal environment size on the number of sperm in (a) the sperm reservoir, (b) the lower isthmus, (c) the upper isthmus, (d) the ampulla, (e) the pre-ampulla and (f) the number of penetrated oocytes. . . . .	149
7.16	(a) The pooled average distribution of sperm for all environments after 8 hours using the calibrated parameters and (b) the average number of capacitated sperm in the ampulla for all environments over time. SR = sperm reservoir, LI = lower isthmus, UI = upper isthmus, A = ampulla, PA = pre-ampulla. . . . .	152
7.17	A comparison between the simulation results and the number of penetrated sperm which are polyspermic from Braden et al. (1954). 'Lit. Table 1' is from Table 1 (pg. 394) and 'Lit. Table 2' is from Table 2 (pg. 395). . . . .	153
7.18	A comparison between the simulation results and data presented by Braden and Austin (1954a). . . . .	154
7.19	A comparison between the simulation results and data presented by McGaughey et al. (1968). . . . .	157
8.1	The impact that (a) individual variations in the 3D environments, (b) variations in the external bends and (c) variations in the internal folds have on the number of sperm in the ampulla. . . . .	162
8.2	The impact that (a) individual variations in the 3D environments, (b) variations in the external bends and (c) variations in the internal folds have on the number of penetrated oocytes. . . . .	163

- 8.3 The impact that (a) individual variations in the 3D environments, (b) variations in the external bends and (c) variations in the internal folds have on the percentage of penetrated oocytes which are polyspermic. . . . 164
- 8.4 The distribution of sperm in (a) the normal environment, (b) the environment without bends, (c) then environment without folds and (d) the environment without bends or folds. SR = sperm reservoir, LI = lower isthmus, UI = upper isthmus, A = ampulla, PA = pre-ampulla. . . . . 165
- 8.5 The relative percentage of capacitated sperm retained in the sperm reservoir under different environmental conditions. Normal = the normal environment, NB = the environment without bends, NF = the environment without folds, NFB = the environment without bends of folds. . . . . 166
- 8.6 The relative percentage of sperm entering from the sperm reservoir which were retained in the lower isthmus under different environmental conditions. Normal = the normal environment, NB = the environment without bends, NF = the environment without folds, NFB = the environment without bends of folds. . . . . 167
- 8.7 The relative percentage of sperm entering from the lower isthmus which were retained in the upper isthmus under different environmental conditions. Normal = the normal environment, NB = the environment without bends, NF = the environment without folds, NFB = the environment without bends of folds. . . . . 167
- 8.8 The relative percentage of sperm entering from the upper isthmus which were retained in the ampulla under different environmental conditions. Normal = the normal environment, NB = the environment without bends, NF = the environment without folds, NFB = the environment without bends of folds. . . . . 168
- 8.9 The relative percentage of sperm entering from the upper isthmus which passed through the ampulla and remained in the pre-ampulla under different environmental conditions. Normal = the normal environment, NB = the environment without bends, NF = the environment without folds, NFB = the environment without bends of folds. . . . . 169

8.10	(a) The percentage of penetrated oocytes at the end of the simulation and (b) how the percentage of penetrated oocytes changes over time for the different environmental conditions. Normal = the normal environment, NB = the environment without bends, NF = the environment without folds, NFB = the environment without bends of folds. . . . .	169
8.11	The percentage of oocytes which are polyspermic out of all penetrated oocytes under different environmental conditions. Normal = the normal environment, NB = the environment without bends, NF = the environment without folds, NFB = the environment without bends of folds. . .	170
8.12	The relative distribution of sperm within the oviduct after 12 hours for the different sperm population sizes. Error bars show SD. . . . .	173
8.13	How the percentage of penetrated oocytes changes over time for the different sperm population sizes. The horizontal red line marks the average percentage of penetrated oocytes after 8 hours for the control sperm population. . . . .	174
8.14	How the percentage of monospermic oocytes out of all oocytes changes over time for the different sperm population sizes. . . . .	175
8.15	How the percentage of polyspermic oocytes out of all penetrated oocytes changes over time for the different sperm population sizes. . . . .	176
A.1	A porcine oviduct scanned using a Micro-CT scanner. (a) the oviduct filled with air scanned without contrasting agent and (b) the oviduct segment scanned with Barium Sulphate solution contrasting agent. . .	205
C.1	(a) OvAS (b) OvNFS (c) OvA1 (d) OvNF1 . . . . .	211
C.2	(a) OvA2 (b) OvNF2 (c) OvA3 (d) OvNF3 . . . . .	212



# Glossary

**Ampulla** Section of the oviduct that connects to the ovary.

**Capacitation** Physical change in sperm to allow for fertilisation.

**CASA** Computer Assisted Semen Analysis

**Chemotaxis** Movement of cells in response to a chemical attractant

**Cilia** Small, hair like structure attached to cells.

**Coitus** Sexual intercourse

**Conception** See fertilisation

**Copulation** Sexual intercourse

**CUDA** Compute Unified Device Architecture

**Embryo** Egg cell after fertilisation and a specific number of cellular divisions have taken place.

**Epithelium** Cellular layer that lines skin and other organs

**Extramural** Segment of the oviduct just outside the uterus

**Fallopian Tube** Term used to describe the Oviduct in humans

**Fertilisation** Fusion of male and female gametes resulting in an embryo

**Follicular** The phase of the reproductive cycle prior to ovulation.

**Gamete** Generic term for both sperm and oocyte

**GPU** Graphics Processing Unit

**HCG** Human Chorionic Gonadotrophin

**Hyperactivation** State of sperm motility where the sperm moves with a high amplitude, non-symmetrical beat pattern. Typically aids fertilisation.

**Immotile** Not able to move.

**Intramural** Segment of the oviduct inside the uterus

**in vitro** Describes a biological experiment carried out in a lab

**in vivo** Describes a biological experiment carried out within a live animal / human

**Isthmus** Section of the oviduct that connects to the uterus.

**IVF** In-Vitro Fertilisation

**Lumen** The internal space within a tissue

**Luteal** The phase of the reproductive cycle after ovulation.

**Monospermy / Monospermic** An oocyte penetrated by a single sperm

**Motile** Capable of independent and spontaneous movement.

**Oestrous** The phase of the reproductive cycle prior to ovulation, adjective form.

**Oestrus** The phase of the reproductive cycle prior to ovulation, event form.

**Oocyte** The egg, release during Ovulation

**Oocyte Cumulus Complex (OCC)**  
Collection of cells released from the ovary along with the Oocyte

**Oviduct** Section of the reproductive tract that connects the uterus to the ovary

**Ovulation** The moment where the Oocyte is released from the ovary

**Peri-Ovulatory** The phase of the reproductive cycle prior to ovulation.

**Peristaltic** Rhythmic, wave like contractions

**Phagocytosis** Chemical process within the body to break down and absorb cells

**Polyspermy / Polyspermic** An oocyte penetrated by multiple sperm

**Porcine** Pig

**Post Coitus** After sexual intercourse

**Post-Ovulatory** The phase of the reproductive cycle after ovulation.

**Thermotaxis** Movement of cells in response to a temperature difference

**Uterus** Part of the reproductive system where implantation and foetus development occur.

**UTJ** Uterotubal Junction. Region of the oviduct that connects to the uterus

**Zona Pellucida** Glycoprotein membrane surrounding the oocyte

## Chapter 1

### Introduction

The mammalian female reproductive system is a highly complex and dynamic environment where multiple internal and external components interact in different ways over time, resulting in the creation of life. Changes in modern lifestyles, combined with a large number of physical, psychological, chemical and environmental factors such as stress, smoking, alcohol, lack of exercise and increased maternal age can have a negative impact on male and female fertility (Sharpe, 2000; Younglai et al., 2005). Understanding the interactions which occur within the female reproductive tract will lead to the development of new fertility treatments and a better understand the progression of diseases.

Investigating the dynamics of the reproductive system through biological experimentation can be difficult, as many experimental procedures are not possible due to ethical considerations. There are certain things which are just not possible to do experimentally, such as performing invasive *in vivo* experiments without influencing the system behaviour. The equipment required to perform certain experiments may not exist, or may not be sensitive enough to generate accurate results (Van Soom et al., 2010). Where an experiment is possible both ethically and practically, it may be prohibitively expensive.

One method which can be used to aid our understanding of biological systems is computational modelling. Computational modelling in biological systems is becoming increasingly common, and a wide range of techniques are employed depending on the type of system and the behaviour being investigated. Once constructed and validated, computational models can allow investigation of aspects of the system which cannot be performed experimentally. When compared to other biological fields such as the cardiac system, computational modelling of the reproductive system is still in its infancy.

## 1.1 Overview of mammalian reproduction

Fertilisation in mammals occurs due to a series of complex processes involving sperm, which are the male reproductive cells and oocytes, which are the female reproductive cells. The processes used by sperm to navigate from the site of deposition to the oocytes at the site of fertilisation in the female reproductive tract are highly complex and poorly understood. Although millions of sperm are deposited into the female reproductive system by the male, only a few hundred to a few thousand reach the oviduct. This massive reduction is due to a combination of physical barriers, chemical processes and individual sperm behaviour. Once within the oviduct, sperm are stored in a sperm reservoir and are gradually released to progress to the site of fertilisation.

This progression is achieved through a combination of attachment and detachment to the oviduct wall and switching between movement patterns. It is important that sufficient sperm reach the site of fertilisation during the short fertility period where the oocyte can be penetrated. However, if too many sperm reach the site of fertilisation at the same time, this can result in polyspermy, which is where multiple sperm penetrate the same oocyte, leading to abnormal development and termination. Understanding how the combination of all factors regulates the supply of spermatozoa to the site of fertilisation can be difficult.

The internal structure of the oviduct is highly complex, containing interweaving folds of soft tissue and external bends which wrap and fold over each other, adding further complexity to the internal structure. It has been proposed that the structure of the oviduct is a key factor in the regulated passage of sperm to the site of fertilisation (Suarez, 1987; Zamboni, 1972).

### 1.1.1 Research objectives

The aim of this thesis is to investigate the complex interactions between the individual sperm behaviour and the structural features of the oviduct. To be able to investigate the impact of the 3D structure alone on sperm progression, a system which allows the physical aspects of the oviduct to be modified without changing other mechanisms is required. This is something which is not possible to do biologically with the current technology. Instead, a computational model of sperm behaviour within the 3D environment of the oviduct was constructed. Key questions that the model will allow us to explore are:

- (a) What is the relative significance of different mechanisms and features in the mam-

malian oviduct?

- (b) Is a model with only the basic sperm navigation mechanisms able to predict the time course of successful sperm - oocyte interactions?
- (c) Does the complex 3D structure of the oviduct regulate the passage of sperm to the site of fertilisation?
- (d) Which aspects of the 3D environment are most significant with respect to the regulation of sperm progression to the site of fertilisation?
- (e) How does the number of sperm within the oviduct affect the number of penetrated oocytes?

In order to address these questions, a number of different tools, components and techniques need to be developed. The main research objectives are as follows:

- 1 Review the biological literature surrounding sperm behaviour within the mammalian oviduct and identify the significant processes and parameters.
- 2 Review the computer science literature surrounding computational modelling of biological systems and identify the most suitable techniques for developing a computational model of sperm behaviour.
- 3 Develop a generic conceptual model of sperm navigation within the oviduct.
- 4 Construct a set of 3D computational models of the oviduct for use within the model.
- 5 Identify a suitable set of computational tools for implementing the model.
- 6 Implement a computational model of sperm behaviour and manage the data.
- 7 Develop a method for presenting the results of the model to biologists.
- 8 Validate the computational model of sperm behaviour against multiple literature sources.
- 9 Use the validated computational model to investigate the research questions.

## 1.2 Thesis structure

This section outlines the structure of the thesis and the contents of each chapter in relation to the research objectives identified above.

**Chapter 1: Introduction** provides an introduction to the thesis. The research area is outlined and the processes involved in the mammalian reproductive system are summarised. The research aims are highlighted and the thesis structure is outlined.

**Chapter 2: Reproductive biology** provides a detailed description of the mammalian reproductive system and the different processes and theories surrounding sperm navigation in the oviduct. The individual components of the reproductive system are identified and species specific variations are discussed. This chapter partially addresses research objective (1).

**Chapter 3: Computational modelling of biological systems** provides an overview of computational biology and computational modelling of biological systems. The different approaches and techniques are described and the use of computational modelling to model biological systems and different processes in the reproductive system are reviewed. This chapter fully addresses research objective (2).

**Chapter 4: Conceptual model of sperm behaviour** identifies the scope for investigation and the techniques used to create a generic model of sperm behaviour. The processes and parameters involved are specified, along with the spatial and temporal scales involved. The dynamics of the system and how the processes and parameters relate to each other are described and the different computational components required to implement the model are identified. Finally, parameter values from literature are identified. This chapter partially addresses research objective (1) and fully addresses research objective (3).

**Chapter 5: Creating a 3D virtual oviduct** describes the construction of a set of 3D computational models of the oviduct. Existing techniques for creating 3D models of biological organs are reviewed and the need for new techniques for this application are identified. The techniques used for model construction and feature extraction from histology images are then described. Optimisation techniques used to speed up image processing are also presented. This chapter fully addresses research objective (4).

**Chapter 6: Computational model implementation** describes the implementation of the agent based model of sperm behaviour. An analysis of existing agent frameworks is performed to identify a suitable framework for model implementation. Implementation of the conceptual model of sperm behaviour presented in Chapter 4 is then described, along with optimisation of different aspects of the model such as collision detection and simulation replicates. This chapter fully addresses research objectives (5), (6) and (7).

**Chapter 7: Calibration and validation** describes model calibration and validation against literature. An overview of model calibration and validation techniques is presented and a set of parameters are identified as targets for calibration. The system is calibrated based on sperm population at the site of fertilisation, and then sensitivity analysis is performed on all parameters. Validation is then performed against sperm penetration and incidence of polyspermy. This chapter fully addresses research objective (8) and partially addresses research objective (9).

**Chapter 8: Experimentation** describes two different sets of experiments to investigate how the oviduct regulates sperm navigation. First, the 3D environment of the oviduct is varied in different ways to identify the relative significance of the different structural features. Finally the total sperm population within the oviduct is varied to determine how large variations influence oocyte penetration and polyspermy. This chapter partially addresses research objective (9).

**Chapter 9: Conclusion** is the concluding chapter which identifies the contributions of the thesis and how the research aims specified in Chapter 1 have been addressed.

### 1.3 Contribution to knowledge

The contributions of this work is summarised below. Where already published, the number refers to the list of publications in Section 1.4.1.

- Literature review of sperm behaviour within the oviduct, and the different aspects which have been modelled computationally [Published in (3)]
- Identification of limitations in biological literature from a modelling perspective [Published in (4)]
- Development of the first conceptual agent based model of sperm behaviour [Published in (4) & (5)]

- Development of techniques for constructing accurately scaled 3D environments from histology images [Published in (1) & (2)]
- Techniques for identifying tubular topology from sparse image datasets [Published, in (1) & (2)]
- Comparison of agent based modelling frameworks with respect to performance with large agent populations
- Use of the FLAME GPU agent based modelling framework to run concurrent simulation replicates [Published in (5)]
- Demonstrated the importance of the use of optimisation techniques for continuous collision detection in agent based simulations [Published in (5)]
- Identified the relative significance of different processes and parameters of sperm navigation within the oviduct
- Presented calibration and validation of the agent based model of mouse sperm behaviour

## 1.4 Publications

Several of the chapters have been partially published. The publication number in brackets relates to the list of peer reviewed publications shown below. A large amount of content from Chapters 2 and 3 has been published in (3). Aspects of Chapter 4 have been published in (4) & (5). Most of Chapter 5 has been published in (1) & (2). The second half of Chapter 6 has been published in (5). Below is a list of all peer reviewed papers, published abstracts and posters.

### 1.4.1 Peer reviewed papers

- (1) Burkitt, M., Walker, D., Romano, D. M. and Fazeli, A. (2010), Using the GPU and Multi-Core CPU to Generate a 3D Oviduct Through Feature Extraction from Histology Slides. *in* ‘2nd International Workshop on High Performance Computational Systems Biology (HiBi 2010)’. University of Twente, Netherlands. Oct. 2010. pp. 78 - 87. [**Conference Proceedings**]
- (2) Burkitt, M., Walker, D., Romano, D. M. and Fazeli, A. (In Press), ‘Constructing complex 3D biological environments from medical imaging using high performance computing.’ *IEEE/ACM Transactions on Computational Biology and Bioinformatics*. [**Journal**]



- (3) Burkitt, M., Walker, D., Romano, D. M. and Fazeli, A. (In Press), ‘Computational modelling in maternal interactions with spermatozoa; potentials and prospects’. *Reproduction, Fertility and Development*. [Journal]
- (4) Burkitt, M., Walker, D., Romano, D. M. and Fazeli, A. (In Press), ‘Using Computational modelling to investigate sperm navigation and transport in the female reproductive tract’. *Theriogenology*. [Journal]
- (5) Burkitt, M., Walker, D., Romano, D. M. and Fazeli, A. (2011), ‘Modelling sperm behaviour in a 3D environment.’ in ‘CMSB 2011, Proceedings of the 9th International Conference on Computational Methods in Systems Biology’. ACM SIG Bioinformatics, Institut Henri Poincaré: Paris, France. Sept. 2011. pp. 141-149. [Conference Proceedings]

#### 1.4.2 Abstracts, presentations & posters

- Burkitt, M., Walker, D., Romano, D. M. and Fazeli, A. (2010), Using Computational Modelling to Understand Sperm Transport and Navigation in the Female Reproductive Tract. in ‘Current issues in Andrology for Health & Drug Development’, AstraZeneca Pharmaceuticals, Alderley Park, Cheshire, SK10 4TG. Annual Meeting British Andrology Society. [Abstract, Poster & Presentation]
- Burkitt, M., Walker, D., Romano, D. M. and Fazeli, A. (2010), Using Computational Systems Biology to Investigate Sperm Navigation and Transport in the Female Reproductive Tract. in A. Fazeli & J. Grizelj (eds.), ‘Systems Biology in Maternal Communication With Gametes and Embryos’. GEMINI Cost Action FA0702. [Abstract, Poster & Presentation]
- Burkitt, M., Walker, D., Romano, D. M., and Fazeli, A. (2011), Computational modelling of sperm movements within the oviduct, in ‘Maternal Interaction With Gametes and Embryos’, Gijon, Spain: GEMINI Cost Action FA0702, Sep. 2011. [Abstract & Poster]

## Chapter 2

# Reproductive biology

This chapter provides an overview of the mammalian female reproductive system. The important individual components are described, along with the complete process of sperm progression from the site of deposition until they meet the oocyte at the site of fertilisation. Different theories of sperm navigation are presented, along with the significance of differences between species and individuals.

### 2.1 Components of fertilisation

In mammals fertilisation occurs within the oviduct, which is called the fallopian tube in humans. The oviduct is a convoluted tube which connects the uterus to the ovaries. Figure 2.1 shows the different regions of the oviduct, and how they are linked to the remainder of the female reproductive tract. The main sections of the oviduct are the isthmus region and the ampulla region. The two main types of cell which are involved in the fertilisation process are oocytes and spermatozoa, referred to collectively as gametes. Spermatozoa advance up the isthmus from the uterus, and oocytes are released from the ovary and progress down the ampulla. Spermatozoa meet the oocytes at the site of fertilisation within the ampulla.

#### 2.1.1 Spermatozoa

Spermatozoa, which are commonly referred to as sperm, are the male reproductive cells in mammals. They are single haploid cells containing 23 chromosomes. Figure 2.2 shows a breakdown of the basic components of a sperm cell. The cell nucleus is contained within the head, and the midpiece contains a mitochondria coil, which provides a source of energy. Sperm cells commonly have a single flagellum or tail, although species specific variations do exist. Spermatozoa are autonomous individuals and move through liquid by synchronous or asynchronous flagellar beats.

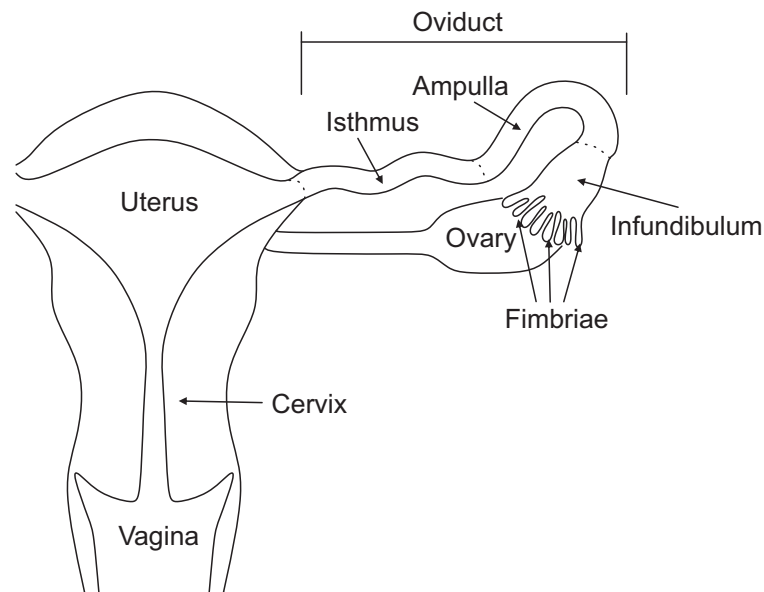


Figure 2.1: The oviduct in relation to the rest of the female reproductive system. The diagram is based on human anatomy.

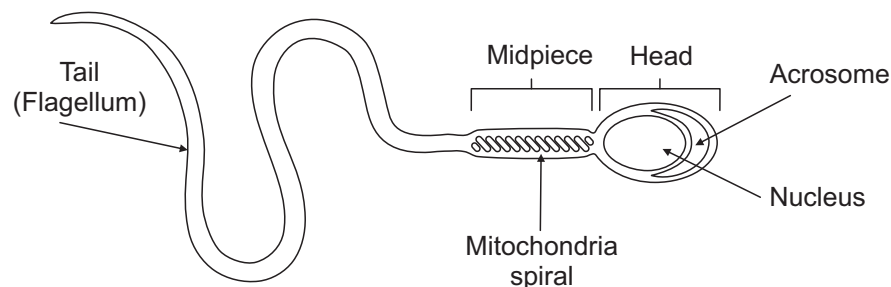


Figure 2.2: A schematic diagram of the basic components of a spermatozoon.

Spermatozoa from different species have different swimming patterns (Phillips, 1972). Despite these differences, most motile spermatozoa have two distinct movement states, which are progressive movement and non-progressive movement. Progressive movement, which is characterised by regular, synchronous flagella movements, results in an almost straight forward movement pattern. Non-progressive movement, which is characterised by high amplitude, rapid, asynchronous flagella beats, results in a highly erratic, irregular movement path, and complex turning behaviour in many species (Goodson et al., 2011; Ho and Suarez, 2001; Phillips, 1972; Suarez et al., 1991). The term *hyperactivation* is frequently used to describe the non-progressive movement pattern (Ho and Suarez, 2001). Non-progressive movement of spermatozoa has been

linked to an increase of calcium ions ( $Ca^{2+}$ ), and results in more energetic swimming patterns (Suarez, 2008a). It has also been suggested that the direction in which spermatozoa move may be determined by modulation between progressive and non-progressive movement patterns (Suarez, 2008a).

### 2.1.2 Oocytes

Oocytes, which are commonly referred to as eggs, are the female reproductive cells in mammals. Like spermatozoa, they are single haploid cells containing 23 chromosomes. Figure 2.3 shows a breakdown of the basic components of an oocyte. Oocytes are surrounded by the zona pellucida, which is a glycoprotein membrane. The zona pellucida is itself surrounded by a few layers of follicular cells called the corona radiata. When released from the ovary, an oocyte is contained within a collection of cells called the oocyte cumulus complex (OCC) (Johnson and Everitt, 1999).

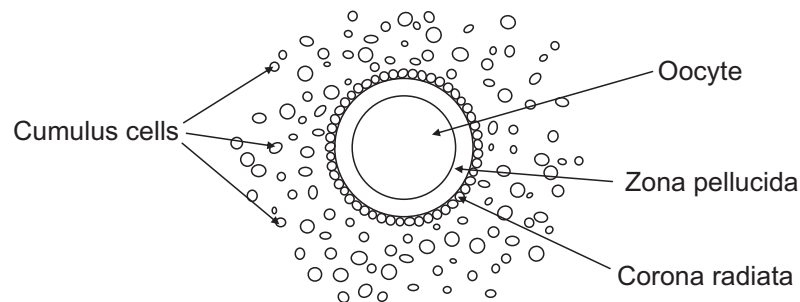


Figure 2.3: A schematic diagram of the basic components of an oocyte with surrounding cumulus complex.

### 2.1.3 The oviduct

The oviduct is often thought of as a simple tube. However, in reality it has a complex internal structure and wraps around itself and other organs in the body. The oviduct is the site of fertilisation, and recent studies have highlighted the importance of interactions between the soft tissue folds of the oviductal walls, the sperm and egg with regards to successful fertilisation, early embryo development, and problems during pregnancy (Fazeli and Pewsey, 2008).

The oviduct tube physically interacts with other organs, changing its shape and position. In addition to this, the internal structure of the oviduct is constantly changing. The muscles in the oviduct wall, in combination with the uterus and ovary, act as a peristaltic pump (Zervomanolakis et al., 2007). This pump causes muscular contractions, which change in relation to the current phase of the oestrous or menstrual cycle,

and are strongest in the peri-ovulatory phase (Wanggren et al., 2008).

The oviductal epithelium is lined with a combination of ciliated cells and secretory cells. The ciliated cells, which have the appearance of small hairs, beat with varying frequency and direction, in relation to hormonal changes and the current phase of the oestrous or menstrual cycle (Chilvers and O’Callaghan, 2000; Lyons et al., 2006). The secretory cells, which change in size, expand as the period of ovulation approaches, and then release their contents at ovulation, reducing in size (Lyons et al., 2006). The proportional composition of ciliated and secretory cells also changes along the length of the oviduct and in relation to hormonal changes (Lyons et al., 2006). When viewed using Scanning Electron Microscopy (SEM), the surface of the oviduct contains crypts, branches and other complex surface structures, which would likely interfere with the passage of sperm (Yaniz et al., 2006).

Within the oviduct, several distinctive regions can be identified. The tube starts with the UTJ, which is a short area of transition connected to the uterus (Hafez and Black, 1969). Following this are the intramural segment of the isthmus, which is embedded within the uterus wall, and the extramural segment, which is just outside the uterus wall. The tube then narrows forming the main section of the isthmus, and makes up approximately one third of the total length (Beck and Boots, 1972). The tube then widens to become the ampulla, and finally the infundibulum and fimbriae.

When looked at from the outside, the oviductal tube has different forms depending on the species. In small mammals such as the mouse, the oviduct tube follows a complex, torturous path which wraps around and folds over itself multiple times (Beck and Boots, 1972). In larger mammals, the structure is very different. In the pig oviduct, the tube is relatively straight, with only a few bends and folds, and the human oviduct is almost entirely straight (Beck and Boots, 1972). Inside the oviduct, there are folds of soft tissue which increase in complexity between the isthmus and the ampulla. In the mouse, these folds are relatively simple, however, in the pig and human, the internal folds are highly complex.

It has been proposed that interactions between swimming sperm and the complex physical structure of the oviduct could be key to the regulation of sperm passage to the site of fertilisation (Suarez, 1987; Zamboni, 1972). The internal and external structures of the individual sections of the tube are important for regulation of sperm movement, and it has been shown that when the isthmus is surgically removed and the remainder of the tube reconnected, a large increase of polyspermy occurs (Hunter and Léglise, 1971).

## 2.2 The reproductive process

Successful fertilisation is achieved through the interaction of two different processes, ovulation and sperm progression.

### 2.2.1 Ovulation

Ovulation marks the point in the reproductive cycle where oocytes are released from the ovary, and the majority of gamete interactions within the oviduct occur shortly after ovulation. Most mammalian females undergo regular cyclic changes to their reproductive system. These changes are referred to as either the menstrual cycle, where the lining of the uterus is shed, or the oestrous cycle, where the lining of the uterus is reabsorbed (Johnson and Everitt, 1999). This cycle can be split into several different phases, and transitions between phases are triggered by hormonal changes. Figures 2.4 shows a generalisation of the different phases of the menstrual and oestrous cycle in humans and animals. The ‘Fertility Period’ relates to the time frame in which coitus can result in successful conception. Outside of this time window, spermatozoa are either prevented from entering the oviduct (Mastroianni Jr, 1999), or are removed by Phagocytosis (Oren-Benaroya et al., 2007) before ovulation.

At the point of ovulation (Figure 2.4), one or more primary oocytes are released from the ovary. They then undergo a meiotic division and become secondary oocytes. The oocytes are moved into the ampulla region of the oviduct by the fimbriae and then move towards the ampullary-isthmic junction, which is the section of the oviduct where the isthmus joins the ampulla. Several different processes may contribute to the movement of oocytes down the oviduct. The beating of cilia, which are hair like appendages to the epithelial cells lining the oviduct, has been identified as a highly significant factor in oocyte transport. Problems with either the beat frequency or the adhesive properties of cilia can impede oocyte transport (Lyons et al., 2006).

### 2.2.2 Sperm progression

After coitus, semen, composed of spermatozoa and seminal plasma, is deposited in the cervix, uterus or vagina, depending on the species (Suarez and Pacey, 2006). Spermatozoa then move through the uterus towards the isthmic region of the oviduct. Entrance of spermatozoa into the oviduct is restricted by the utero-tubal junction (UTJ), which helps to prevent immotile or damaged spermatozoa and foreign particles from entering the site of fertilisation (Dukelow and Riegler, 1972). The thick mucus and contracted muscle in the isthmus may also prevent spermatozoa from passing through quickly.

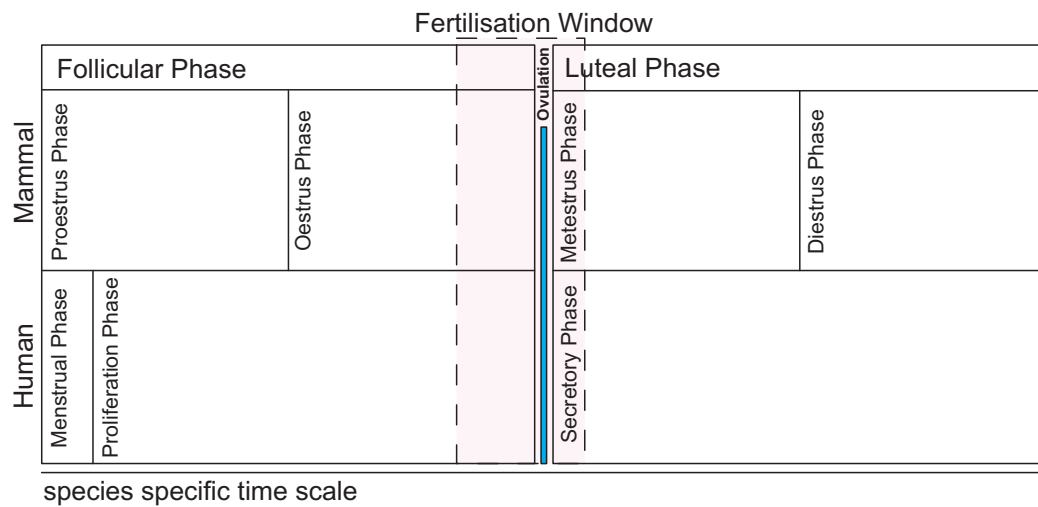


Figure 2.4: The different phases of the menstrual cycle in humans and the oestrous cycle in mammals. The ‘Fertility Period’ relates to the time frame in which coitus can result in successful conception. Based on descriptions presented in Johnson and Everitt (1999).

The number of spermatozoa that arrive in the oviduct is several degrees of magnitude smaller than the number initially deposited (Suarez and Pacey, 2006). However, despite large variations in the number of spermatozoa deposited by different species, a similar number of spermatozoa reach the oviduct. (Harper, 1982)

The common belief that large numbers of spermatozoa race each other to reach and fertilise the oocyte is no longer considered to be correct (Eisenbach and Giojalas, 2006; Foo and Lim, 2008), and more complex guidance mechanisms are now thought to control the process. In many mammals, it has been observed that a ‘sperm reservoir’ is established either outside the oviduct or within the caudal isthmus (Hunter, 1981; Smith and Yanagimachi, 1991; Suarez, 2002). This is achieved by either spermatozoa binding to the oviductal epithelium due to secretions from the oviduct that suppress motility (Suarez, 1987), or becoming trapped within the thick isthmic mucus and soft tissue folds (Suarez and Pacey, 2006).

The purpose of this reservoir appears to be to allow storage of spermatozoa until the ovulation (Suarez, 2002). Attachment to the oviductal epithelium has been shown to provide a nutrient rich environment for spermatozoa, thus prolonging their motility (Holt and Fazeli, 2010; Suarez, 2008b), although this does not extend the life of spermatozoa beyond their maximum lifetime. Spermatozoa stored in the sperm reservoir are gradually released from their epithelial bindings, a process which has been linked to capacitation (Hunter, 2008; Lefebvre and Suarez, 1996; Suarez, 2002).

Capacitation is a maturation process that spermatozoa undergo to allow them to penetrate the oocyte (Bedford, 1983; Jaiswal and Eisenbach, 2002). In some species, capacitation has been observed to cause changes in the proteins lining the sperm head surface, resulting in a reduction in the number of spermatozoa bound to epithelial tissue (Lefebvre and Suarez, 1996; Suarez, 2002). The detachment of capacitated spermatozoa from the oviduct epithelium is further facilitated by their own vigorous movement patterns (Demott and Suarez, 1992; Lefebvre and Suarez, 1996).

Only a small percentage of the sperm population is capacitated and active at any one time (Giojalas et al., 2004; Tienthai et al., 2004). Once capacitated, spermatozoa navigate the complex internal environment, gradually detaching, moving and then reattaching to the epithelium. Spermatozoa do not attach every time they touch epithelium, and sometimes swim off in a different direction (Demott and Suarez, 1992). Capacitated spermatozoa only live for a short period of time, and if they do not encounter the oocyte soon after capacitation they will die (Bedford, 1983; Cohen-Dayag et al., 1995). This results in a gradual release of spermatozoa over time and provides a stream of spermatozoa to the site of fertilisation. Eventually, a small number of spermatozoa reach and attach to the oocytes in the ampulla and the fertilisation process begins.

Fertilisation itself is a complex process (Ikawa et al., 2010). At a basic level, sperm that are attached to the oocyte rapidly penetrate the Zona Pellucida and then fuse to the inner cytoplasmic mass (Austin and Braden, 1956; Jin et al., 2011). If multiple sperm penetrate the oocyte, then polyspermy can occur. Polyspermy is where more than one sperm fertilises the oocyte, and results in abnormal development and termination. Once penetrated, oocytes undergo a reaction which prevents further sperm from penetrating the egg membrane, thus establishing a block to polyspermy (Wolf, 1978).

### 2.2.3 Complex guidance mechanisms

Several complex mechanisms, such as the movement of oviductal fluid (Fauci and Dillon, 2006); oviductal contractions (Zervomanolakis et al., 2007); local communication between gametes and the oviduct epithelium (Seytanoglu et al., 2008; Sostaric et al., 2006); and sperm response to temperature and chemical concentration gradient (Eisenbach and Giojalas, 2006) within the oviduct have been investigated to determine their importance with regard to sperm transport to the site of fertilisation.



### 2.2.3.1 Fluid dynamics

The lumen of the oviduct contains fluid of varying viscosity and chemical composition (Suarez and Pacey, 2006). It has been proposed that understanding the behaviour of fluids within the oviduct is key to understanding the reproductive process (Fauci and Dillon, 2006). Unlike blood, which moves through arteries in a single direction, the fluid within the oviduct is more dynamic. The composition of different substances within the fluid is constantly changing, and fluids with differing viscosity and shear forces could influence the overall flow (Brokaw, 2001; Riffell and Zimmer, 2007). This is especially true at the time of ovulation, where secretory cells within the oviductal lumen release their contents. In rodents, the sudden release of liquid combined with the presence of oocytes in the ampulla, result in the walls of the ampulla becoming distended and stretched (Nilsson and Reinius, 1969; Zamboni, 1972).

Around the time of ovulation, the walls of the uterus and the oviduct undergo peristaltic contractions. Due to these contractions, and the consequential movement of fluid towards the site of fertilisation, it has been shown that spermatozoa are present in the oviduct only a few minutes after copulation (Suarez and Pacey, 2006). However, these spermatozoa have been found to be damaged by the forces applied to them, rendering them immotile (Suarez and Pacey, 2006). In the hamster, particles of Indian ink injected into the isthmus of the oviduct are quickly transported to the site of fertilisation during the peri-ovulatory period, due to the coordinated contractions in both the isthmus and the ampulla (Battalia and Yanagimachi, 1979).

By contrast, it has been shown that when flushing the oviduct, the initial flushing contains mainly immotile spermatozoa, with the most motile spermatozoa being strongly attached to the epithelial surface, and more vigorous flushing methods are required to detach these motile spermatozoa (Smith and Yanagimachi, 1990). This implies that fluid movement alone is not sufficient to move the most motile spermatozoa to the ampulla, and a combination of other mechanisms is required. Also, studies of spermatozoa *in vitro* have shown that a subset of spermatozoa actively swim against a fluid flow (Seo et al., 2007).

Fluid dynamics may have some influence on sperm transport, but it is unlikely to be the only factor controlling sperm movements. Oviductal contractions (Zervomanolakis et al., 2007), cilia beating (Fauci and Dillon, 2006; Lyons et al., 2006) and the structure of the folds within the oviduct could also impact the movement of fluid and spermatozoa, but it is more likely to result in mixing of the luminal contents than a controlled transport of spermatozoa (Harper, 1982; Perkins, 1972).

### 2.2.3.2 Local maternal communication

‘Maternal communication’ between the oviduct, gametes and embryos has been the focus of a number of recent investigations (Fazeli and Pewsey, 2008). The presence of gametes within the oviduct appears to alter the local chemical composition of excretions produced by the oviductal epithelia, which in turn alters gamete behaviour and can have major influence on sperm movement and transport within oviduct (Georgiou et al., 2007; Seytanoglu et al., 2008). This creates a dynamic environment for gametes and embryos to interact with the female reproductive tract.

### 2.2.3.3 Thermotaxis & chemotaxis

It has been suggested that spermatozoa use both chemotaxis and thermotaxis as guidance mechanisms, with thermotaxis acting as a long range attractant and chemotaxis as a short range attractant (Eisenbach and Giojalas, 2006).

At the time of ovulation, the temperature in the isthmic region has been shown *in vivo* to drop in some mammals (Bahat et al., 2003), thus establishing a temperature gradient between the cooler isthmic and warmer ampulla regions of oviduct. Spermatozoa from rabbits and humans have been shown *in vitro* to follow an artificial temperature gradient from cool to warm (Bahat et al., 2003). However, this does not demonstrate that spermatozoa are able to detect the relatively shallow temperature gradient which may exist in a real oviduct (Eisenbach and Giojalas, 2006). A difference in temperature between two different regions does not guarantee that a constant, gradual change in temperature will be established along the length of entire tube, and it is likely to be evenly distributed across the organ tissue. This means that in species where the oviduct is relatively straight, such as the pig or human, this theory is viable. However, in species where the tube wraps around itself, such as in rodents, this theory is less viable.

Spermatozoa from different species have been shown to be attracted to chemicals released by the OCC and the oocyte. The actual chemical that attracts spermatozoon has not been identified, and the possibility of multiple attractants has been suggested (Eisenbach and Giojalas, 2006). Spermatozoa in humans have been shown *in vitro* to follow an attractant to the OCC, then to follow a different attractant to the oocyte (Sun et al., 2005). The attraction of sperm to chemicals released from the OCC *in vitro* does not guarantee that this is a highly significant factor *in vivo*. The movement of fluid within the lumen, secretions from the oviductal epithelium and oviductal contractions are likely to disrupt any chemical gradients within the oviduct.

### 2.3 Species specific variations

When viewed from a high level, the generic behaviour of the mammalian reproductive system is similar between different species, in that spermatozoa, deposited by the male at copulation, move through the female reproductive tract to encounter the oocyte within the ampulla of the oviduct for fertilisation. When looked at in detail however, significant differences are observable in different species, the most apparent of which is the difference in scale. For example, the oviduct of the mouse is very small in comparison to larger mammals such as pigs or humans. When examined in detail, the structure of the oviductal tube is also different between species. The mouse oviduct is a convoluted tube which wraps around and folds over itself multiple times (Beck and Boots, 1972). In larger mammals, the structure is very different. In the pig oviduct, the tube is relatively straight, with only a few bends and folds, and the human oviduct is almost entirely straight (Beck and Boots, 1972).

In many mammals, a sperm reservoir is established in the intramural and extramural segments of the isthmus by attachment of the sperm head to the epithelial tissue (Suarez and Pacey, 2006). In humans, this does not happen in the same way (Williams et al., 1993), and it is postulated that the sperm reservoir is established purely due to the complexity of the internal folds and thick mucus restricting movement of sperm (Suarez and Pacey, 2006).

Significant differences in the size of the sperm population also exist. In mice, around 50 million spermatozoa are ejaculated, where as in pigs the figure is around 8,000 million (Harper, 1982). However, it has been noted that despite these massive differences, a similar number of spermatozoa are found within the oviduct in many species (Harper, 1982). In humans it is usual to have only 1 oocyte at ovulation, and having 2 or more, which may result in twins, is relatively rare. In mice, it is common to have multiple eggs and multiple offspring, with typically 4 oocytes per oviduct at ovulation (Braden and Austin, 1954a).

The behaviour of the individual cells is also different. The size, speed and behaviour of spermatozoa from different species vary greatly. Mouse spermatozoa have flat heads with hooks on the end, an average head width of around  $3.6\mu m$ ; an average total length of around  $123.6\mu m$  (Cummins and Woodall, 1985; Firman and Simmons, 2010); and an average progressive movement speed of around  $147\mu m s^{-1}$  (Goodson et al., 2011). In contrast, human spermatozoa have elliptical heads; have an average head width of around  $3.2\mu m$ ; an average total length of around  $58.5\mu m$  (Cummins and Woodall, 1985) and an average progressive movement speed of around  $60\mu m s^{-1}$  (Mazz-

illi et al., 2001). The difference in morphology could account for the different movement speeds (Gil et al., 2009; Humphries et al., 2008; Malo et al., 2006).

Given the differences in scale between the human and mouse, the fact that mouse spermatozoa swim significantly faster than human spermatozoa is surprising. This difference may be compensated for by differences in oviduct morphology and differences in the lifetime of spermatozoa between species. In mice, spermatozoa have been observed to live for 12 to 24 hours (Benitz and Graves, 1972; McGaughey et al., 1968; Smith and Lodge, 1987), whereas in humans this figure is longer, between 24 to 48 hours (Harper, 1982).

Significant differences also exist within the same species. Different strains of mice have a different number of oocytes per oviduct, a different number of sperm, with different morphology and movement characteristics (Braden, 1958, 1959). With so much variation in the structure and behaviour within different species, understanding the processes can be difficult. This is especially true when considering the possibility that different mechanisms may play more or less prominent roles in the reproductive process depending on the species. Literature reviews of the reproductive process commonly describe the general process based on information from different species (Eisenbach and Giojalas, 2006; Ikawa et al., 2010; Suarez, 2008b), although a few do focus on individual species (Foo and Lim, 2008; Suarez and Pacey, 2006; Zamboni, 1972).

## 2.4 Conclusions

This chapter has outlined the components and processes within the mammalian female reproductive system and how their combined interaction results in the correct number of sperm reaching oocytes at the site of fertilisation at the correct time. Different complementary theories of sperm transport and navigation have been proposed, and the subset of theories which will be investigated have been identified.

Investigating sperm movement within the oviduct is an interesting problem. There is *in vitro* evidence for many of the individual mechanisms involved, however little is known about how those mechanisms work together *in vivo* and how the complexity of the oviductal environment influences them. The focus of this thesis is to create a computational model of sperm behaviour to investigate the significance of the 3D environment in regulating the passage of sperm to the site of fertilisation.

Chapter 3 describes the techniques commonly used for computational modelling of biological systems and reviews the modelling work previously carried out in relation to the reproductive system. Chapter 4 presents the conceptual model, in which the

aspects of the system that will be investigated are identified and the agent model is defined. Chapter 5 describes the creation of the 3D virtual oviduct and Chapter 6 describes the computational implementation of the agent model.

## **Chapter 3**

# **Computational modelling of biological systems**

This chapter provides a description of computational modelling of biological systems and associated techniques. The benefits of computational modelling are highlighted, and the commonly used techniques for modelling different types of biological system are described. The relative work in computational modelling of the reproductive system is then described.

### **3.1 The complexity of biological systems**

Biological entities are not simple structures and computational modelling of biological systems requires the representation of different scales. Biological organs can be broken down into a hierarchy of interconnected subsystems and components. At the top level, an organ interacts with other organs within the individual and the system as a whole. An organ is composed of different types of tissue, each with a specific function or set of functions. Each tissue is made up from collections of cells with similar variations in function. Each cell, which is considered to be the ‘structural and functional unit’ of most forms of life, is made up from several different proteins. A protein is made up of chains of amino acids, the configuration and structure of which determine the type and function of the protein. The most basic component, the gene, is a section of DNA which performs a specific function. Genes are involved in low level operations, such as protein and mRNA creation. Clearly, biological systems are highly complex, and can be modelled at very different spatial and temporal scales. Because of this, several different approaches are used when modelling complex biological systems.

### **3.2 Computational biology**

The term computational biology encompasses many related and in part overlapping areas such as bioinformatics, mathematical modelling and computational modelling

and simulation.

Bioinformatics focuses on the analysis of massive data sets, usually obtained from high-throughput sequencing, using techniques ranging from statistical analysis, data mining, pattern recognition and algorithmic analysis to understand data sets which are too large to analyse manually (Pop and Salzberg, 2008). Bioinformatics is commonly used in molecular cell biology to analyse genomic and proteomic data (Wilke, 2003). Bioinformatic analysis and related techniques are useful for identifying low level processes, such as how gene expression within cells or the protein excretion from cells change in response to different experimental conditions. However, it can be difficult to understand how these low level changes influence the system at the tissue or organ level.

Mathematical modelling looks at biological systems from a different perspective, by using mathematical equations to describe biological processes (Murray, 2007). Mathematical models can take many different forms, depending on the type of system being modelled. They can be statistical, where relationships between the experimental variables are identified and quantified from analysis of experimental results (Draghici et al., 2007; Jeong et al., 2000). They can also be investigative, where mathematical descriptions of real world mechanisms are postulated, and the equations solved or approximated analytically (Anderson et al., 2000; McDougall et al., 2006). These types of model can be useful for identifying linear relationships, but as the number of variables and their relationships become more complex and non-linear, the models become increasingly difficult to understand (Foote, 2007).

Computational modelling and simulation is a way of investigating problems that are too complex to easily solve analytically. A set of parameters and mechanisms which describe how different components of a complex system relate to each other are identified. This is typically based on the laws of physics or known or postulated behaviours where possible, with specific parameters and rules guided by empirical evidence. By varying the input parameters, the differences in output can be observed and used to propose hypotheses on how the underlying components of the system are behaving.

### **3.2.1 Computational systems biology**

Computational systems biology is concerned with integrating large datasets using complex data structures and algorithms to accurately recreate the behaviour of a biological system. It is used to help determine how the internal components of complex biological systems interact with each other and with external systems (Kitano, 2002). By investigating the structure of the components that a system is comprised of; how the

components function and how they interact with each other over time; and how this functionality is influenced by varying conditions, a model can be constructed, based on design principles, to recreate this functionality (Kitano, 2002). This allows specific aspects of a complex system to be investigated without the need to accurately reproduce the complete system in fine detail.

Computational systems biology approaches have been applied to many biological systems at a variety of spatial and temporal scales to complement and enhance experimental research (Caiazzo et al., 2011; Southern et al., 2008). Drug discovery has been enhanced by integrating ‘omics’ data, such as genomic, proteomic and metabolomic data, with modelling to help identify new drug targets, leading to a better understanding of the nature of diseases and how the body will react to new drug combinations (Butcher et al., 2004; Hornberg et al., 2006). There are also attempts to identify measurable differences in bio markers in the blood that may indicate disease related changes in protein and gene regulatory networks, with the aim of predicting and preventing the onset of disease in individuals (Hood et al., 2004).

### 3.3 Computational modelling of biological systems

Computational models allow us to investigate aspects of a biological system which are not financially, practically or ethically possible to look at using experimental techniques and are invaluable tools to complement experimentation or to carry out virtual experiments. When used in combination with experimental methods, this can provide a comprehensive view of the system which would be unavailable using a single method.

Computational modelling can help biologists to think about the system they are investigating in different ways and highlight the limitations of their assumptions and experiments. This could lead to more robust experimental designs and more realistic interpretations of data. It can help to identify gaps in knowledge and the significance of these gaps. For some types of computational model, the behaviour can be displayed visually, demonstrating how the simulation progresses over time. This can help to explain biological interactions to non-specialists, as visualisation of behaviour is far easier to understand than a series of graphs and charts. It is important to note that computational models are only as reliable as the information used to construct them. The emergence of behaviour within a model does not guarantee that the phenomenon also occurs within the real system. Interpretation of results generated by a model must be performed with these caveats in mind.



### 3.3.1 Modelling complex systems

Fully understanding every aspect of a complex system can involve too much information for the human mind to process. Computational modelling helps us to investigate complex systems in detail without the need to fully understand every aspect of the system. Computational models can generally be split into three different categories, descriptive models, conceptual models and complex models (Wissel, 1992). Descriptive, or statistical models, show trends, patterns and relationships between data, but do not provide any understanding about the underlying mechanisms (Grimm, 1994).

Complex models are as comprehensive as possible, including all available data on every aspect. However, if a large number of independent parameters are used in a model, it can be difficult, if not impossible to infer any meaningful relationship between variations in input and the changing behaviour observed in output (Grimm, 1994). Obtaining quantitative measurements from biological experimentation can also be difficult, as many techniques only give qualitative data. For example, when performing the experiments described in Appendix B, the stage of estrous that each mouse was in was determined by looking at vaginal swabs under the microscope and visually classifying the number of cells visible using multiple (+) symbols. A single (+) indicating a small number of cells and (+++++) indicating a large number of cells. Another example is when classifying the behaviour of sperm within the mouse oviduct, Demott and Suarez (1992) who reported that “Sometimes the sperm would reattach the first time it contacted the epithelium, and sometimes it would bounce off, frequently making a direction change, and then continue swimming freely”. This statement describes the behaviour, but does not give any quantitative measures, such as the frequency of attachment and reflection or the angles of direction change.

For a model to be useful, understanding the model itself, and the relationships within the model is essential (Grimm, 1999). Conceptual models only contain the important factors required to understand the behaviour being investigated (Wissel, 1992). Although conceptual models do not contain the same level of detail as complex models, they are easier to manage. The relationship between the small set of input parameters and the generated outputs can be identified, allowing every aspect of the model to be explored (Grimm, 1994).

### 3.3.2 Modelling approaches

Several different approaches are commonly adopted when modelling biological systems. The reductionist approach, which is often referred to as the ‘bottom up’ approach, views

the system as a collection of low-level components. The behaviour of each component is individually investigated, and hypotheses of how they interact are formulated. A simulation of the complete system is then constructed based on the hypotheses of how these components interact with each other (Scott, 2004). This approach has been successful in helping to identify individual components of biological systems and has the strong advantage that it can be based on large amounts of real data. However, modelling based on component interactions alone does not always result in the collective behaviour which can be observed in the system being investigated (Sauer et al., 2007).

The success of the reductionist approach relies heavily on the completeness and accuracy of the data used to create the composite components. Using this approach to simulate large, complex systems is currently not feasible, due to a lack of accurate data and the limitations of computational power. The reductionist approach also fails to acknowledge the influence of environmental conditions and other external factors, which could influence how the system responds to different stimuli (Noble, 2002*b*). Instead of modelling the entire system, a subset of data can be used to create a less accurate model. Although not completely accurate, it could still be sufficient to investigate the required behaviour (Box and Draper, 1986; Noble, 2002*b*).

An alternative to the reductionist approach is the ‘top down’ approach. The system is viewed as a collection of high-level entities or ‘black boxes’, with relationships between the inputs and outputs of each entity defined as a set of rules or mathematical equations based on assumptions about the inner workings and observations of system behaviour (Katagiri, 2003). The model can then be progressively refined, breaking up high level entities into collections of lower level entities. This approach is useful for creating a high-level model of a system, and allows the relationships between high-level processes to be investigated. However, it can lead to an over simplification, and has had little success in accurately determining the underlying components of a system (Katagiri, 2003).

By understanding interactions between different layers within a biological system, robust models of biological systems can be created that will react in the same way as a real organism when exposed to different external factors and stimuli. This goal is difficult to achieve using the reductionist approach or the top down approach alone. However, if both these are used to model the same system concurrently, then they can potentially be connected in the middle, providing co-validation of each other (Katagiri, 2003).

Another approach, called integrative or ‘middle out’ modelling, attempts to combine these two approaches in a different way. Instead of starting in at the largest scale,

such as the organism level or smallest scale, such as the molecular level, the system is modelled midway, from the cellular or tissue level (Noble, 2003). Only a subset of the behaviour of a system is modelled and the level of detail and abstraction used for each component is determined by the type of system and the behaviour that is to be investigated (Noble, 2002*b*). By using this modelling approach, the spatial hierarchy can be traversed upwards to confirm that our collective behaviour matches that observed at a system level. The spatial hierarchy can also be traversed downwards to investigate individual processes. This would be necessary when the process is not well defined, or when small changes in the model parameters relating to the process result in large changes in behaviour. The ability to traverse the model hierarchy gives modellers a large amount of flexibility while developing the model. For biological systems, it has been suggested that the cell is the ideal starting point for middle out modelling, as the internal molecular interactions can be aggregated, and a large amount of experimental data is available from a cellular level (Noble, 2006; Walker and Southgate, 2009).

One integrative method, called ‘synthetic microanalysis’ (Auyang, 1998), looks at a system from both the reductionist approach and the ‘top down’ approach. The method is split into two phases, the synthesis phase and the microanalysis phase. In the synthesis phase, a top level view is used to identify the boundaries of the system, and any interactions that exist with external systems. System level properties are then identified, which may be the size of the individual population, or statistical values based on the state of individuals within the system. Individuals and their important characteristics are then identified using information from literature or small experiments. In the microanalysis phase, the rules and relationships of individual are defined to recreate system level behaviour. This approach provides a way of linking the behaviour of individuals with the behaviour observed at a system level.

### 3.3.3 Modelling techniques

The techniques used to model biological systems can be broadly split into two categories, discrete modelling and continuous modelling.

#### 3.3.3.1 Discrete modelling

A discrete model typically represents a system as a collection of discrete entities - either individuals or spatial regions. A simple technique used for discrete modelling is called cellular automata (Wolfram, 1984). A system is represented as a set of ‘cells’ in a two or three dimensional lattice or grid. In this case, a ‘cell’ represents a fixed region in space, rather than a single biological cell. The structure and size of the grid may vary

depending on the system being modelled. Both space and time are discrete, and the state of each cell, which typically relates to a collection of entities, is updated at fixed intervals based on the state of its neighbours and the environment. Although useful for modelling a variety of different problems, cellular automata can be too simplistic and limiting for modelling more complex behaviours.

Agent based modelling (ABM), which is also referred to as individual based modelling (IBM), is a technique used to model the interactions between populations of entities within a system at a local level (Bonabeau, 2002). Each entity is modelled individually and the overall behaviour of the population is an emergent property of interactions between individuals. ABM has been used to model a wide range of phenomenon such as crowd behaviour (Henein and White, 2005; Shendarkar et al., 2006), Social interactions and traffic (Bonabeau, 2002; Gilbert and Terna, 2000), the global economy (Farmer and Foley, 2009), virus progression during a pandemic (Bobashev et al., 2007; Epstein, 2009) and animal population dynamics (Topping et al., 2003). ABM has also been used to model biological entities within individuals. Depending on the biological level being investigated, ABMs can be viewed from either the reductionist approach, or the integrative approach. The use of ABM in biology is described in more detail in Section 3.4.1.

One approach for developing ABMs is to use pattern oriented modelling (Grimm, 1994). For a given system, a naturally occurring pattern of behaviour is identified and the rules and relationships between individuals within the system are defined so that the pattern of behaviour is recreated. Implementing a pattern also forces accurate temporal and spatial scales to be used. The rules and behaviour of individuals can then be modified to investigate the problem being addressed. This results in a realistic model with fewer parameters, making it easier to understand (Grimm et al., 1996). However, there is no guarantee that the rules applied to the individuals within the system are the same as the mechanisms which exist in reality (Grimm et al., 1996).

### 3.3.3.2 Continuous modelling

Continuous models represent the behaviour of a system as a set of mathematical equations. Ordinary differential equations (ODEs) are used when a function is evaluated with respect to a single value, which in modelling is usually time (Zwillinger, 1997). Partial differential equations (PDEs) are used when a function is evaluated with respect to multiple values, which are usually space and time (Zwillinger, 1997). The important, measurable aspects of a system are parameterised, and the mathematical relationship between the parameters determines the equations. The latter can be based

on statistical analysis and measurable observations of a system (a ‘top down’ approach) or postulated behaviour and relationships between the properties of the system (Sauro et al., 2006).

A common use for PDEs in computational modelling is to determine a scalar value for a specific location at a specific moment in time. One example of this is the diffusion of a biochemical substance, where the temporal variation in the concentration of, for instance, a particular hormone or growth factor is required at specific locations. Smith et al. (2002) used PDEs to simulate the movement of RAN, which is involved in the transport of proteins across the nuclear membrane. The model helped to identify the important regulatory factors involved in establishing the steady-state flux across the membrane.

Simple methods for evaluating PDEs are adequate when the environment geometry being modelling is relatively simple. However, when modelling realistic systems, the environment can be too complex, requiring more sophisticated numerical methods instead (Ŝolín, 2005).

One method commonly used to model continuous phenomena in complex geometries is finite element modelling (FEM) (Bathe, 2007). This is a technique where a complex environment is broken down into small spatial regions or ‘elements’ and the value at a given point is then estimated based on interpolation between neighbouring regions.

FEM has been used for different domains, such as organ modelling in surgical planning and simulation (Brock et al., 2005; Tillier et al., 2003), modelling flow and pressure of blood in the arteries (Oshima et al., 2001; Vignon-Clementel et al., 2006) and modelling lung deformation during respiration (Werner et al., 2009).

Discrete modelling techniques are commonly used when looking at systems where the actions or properties of individuals within a population are important. Continuous modelling is commonly used when the average system behaviour is important, and differences and interactions between individual elements within the system are not significant.

### 3.4 Computational modelling applied to biological systems

Computational modelling has been applied to many biological systems at a variety of spatial and temporal scales to complement and enhance experimental research. One area which has received a large amount of work is modelling the cardiac system, and several detailed models of the heart have been constructed (Noble, 2002*a*) (see (Clayton and Panfilov, 2008) for a recent review). These models look at electrical activity within

the heart from a cellular level, sub-cellular level, tissue level and whole organ level, with the aim of simulating the effects of heart disease, cardiac arrhythmia and other cardiac related afflictions.

Computational simulations of the flow of fluid or gas within 3D models of biological organs are now commonplace and are used for a variety of applications such as; identifying organ damage due to changes in flow rate or pressure (Debbaut et al., 2011); assessing the impact of structural changes in the organ on fluid movement (Chaichana et al., 2011); and the investigating the risk of rupture due to tissue damage such as an aneurysm (Marzo et al., 2011).

Other examples of applications of computational modelling are; studying social insect networks (Jackson et al., 2004); understanding the course of complex diseases such as cancer (Hornberg et al., 2006) and arthritis (Attur et al., 2002); and identifying genes that are linked to specific traits in plant growth (Salekdeh and Komatsu, 2007).

### 3.4.1 Agent based modelling of cell behaviour

There has recently been an increase in the use of ABM techniques to model the behaviour of cells and molecules within the body (Thorne et al., 2007).

Walker et al. (2004) used ABM techniques to model individual epithelial cells in combination with *in vitro* modelling to investigate the influence of  $[Ca^{2+}]$  on wound healing in epithelial cell monolayers. Epithelial cells were modelled individually and given a simple set of rules. These rules specified how the cells should behave in relation to their contact with other cells and their environment. The model was used to formulate predictions about cell proliferation in response to narrow wounds. This prediction was confirmed *in vitro*, showing that a simple set of rules is sufficient to form qualitative predictions about complex mechanisms, such as the influence of calcium on wound closure.

d’Inverno and Saunders (2005) demonstrated the advantages of using ABM techniques over cellular automata to model the self-organisation of stem cells. The agent model was better able to simulate the reversibility and plasticity of cell types.

Different aspects of the immune system have been modelled. Bailey et al. (2007) modelled the movement of inflammatory cells within blood vessels in the muscle in 2D. The agent model was coupled to a network flow model which modelled the flow of blood. Chemokine and cytokine concentrations were modelled using a set of plasma agents, which maintain concentrations of mediators expressed by adjacent inflammatory cells, and degraded over time. Rolling, adhesion and migration behaviours were

implemented based on rules from literature. The model was validated by knocking out adhesion molecule expression, and comparing the results with that reported in literature. An extended version of this model was then used to investigate stromal cell trafficking in Ischemia, leading to a new understanding of the adipose-derived stromal cells (Bailey et al., 2009).

Segovia-Juarez et al. (2004) used ABM to model the formation of Granuloma during *Mycobacterium tuberculosis* infection in 2D. The behaviour of individual immunity cells and bacteria were modelled and simple single variable diffusion was used to model chemokine concentrations within the environment. The model indicated that lower bacterial growth rates actually enhanced survival.

Flugge et al. (2009) used ABM to model the formation of Granuloma as an immune response within the Liver. Different cell types and individual proteins were modelled as agents. Unlike other models of immunity which use 2D grids, Flugge et al. (2009) used a 3D representation of the sinusoids of the liver, inside which the agents interacted. Several hypotheses were tested, and Granulomas were formed in all. They found that the use of a 3D model helped them to think about how the agents interact within the real 3D environment, and aided understanding with biologists when presented as a 3D visualisation.

Mi et al. (2007) developed a model to investigate inflammation and wound healing for diabetic foot ulcers in 2D. The model was able to predict how changes in mediator concentrations effect the inflammation and wound healing process.

The previously described models of immunity are closely linked to literature and biological experimentation, with parameters and mechanisms linked to multiple sources. Jacob et al. (2004) have developed a more simplistic, swarm based 3D model of the immune system. The behaviour of a large number of immunity cells is implemented based on qualitative rules. However, it is unclear how the value of parameters is obtained and no validation is provided.

Cancer is another area where ABM is being used. The behaviour of cancer cells in brain tumours was investigated by Mansury et al. (2002). Individual cancer cells were modelled in a 2D grid using ABM, with rules defining proliferation and migration based on concentrations of nutrients and toxic metabolites within the environment. Depending on the search capabilities of the cells, whether they search globally or locally, different migration and clustering patterns were found. This model was later linked to a sub-cellular network model of the epidermal growth factor receptor (EGFR) signalling pathway to investigate the conditions under which cells switch between proliferation

and migration (Athale et al., 2005). Wang et al. (2007) developed a multi-scale model of tumour expansion in the lungs in 2D. This was later extended to investigate the effects of specific mediators, namely Tumour Growth Factor Beta and Epidermal Growth Factor, on tumour growth in a 3D block of tissue (Wang et al., 2009).

Clearly, the use of ABM to model individual cell behaviour in a spatial environment is a good approach which has been successfully applied to other systems.

### 3.4.2 Modelling in reproductive biology and related systems

Computational modelling in the field of reproductive biology and medicine is relatively young and the majority of computational models in this field are related to embryo development, viewed from the molecular level. Much of this work has been carried out on *Drosophila*, where the embryo has a very well defined structure and the gene regulatory networks are relatively well characterised (Segal et al., 2008; Zinzen et al., 2006). Computational modelling in the *Drosophila* embryo has also been used to predict localised changes in gene expression in the developing embryo with respect to time (Samsonova et al., 2007).

Although there is an absence of existing literature on computational modelling within the oviduct at a cellular level, there has been computational modelling work performed on other cellular level processes which are related to the oviduct. The oviduct epithelia is lined with ciliated cells, which are involved in spermatozoa and oocyte transport (Lyons et al., 2006). Spermatozoa move within the oviduct using flagella for propulsion within viscous fluid. Low level flagella and cilia behaviour, and the influence of viscous fluid on cellular movement have all been investigated in detail and different computational models have been created in the past. Cilia, which are the small, hair like cellular appendages which can be found in many biological structures such as the respiratory, digestive and reproductive systems, have been investigated in depth (Brokaw and Luck, 1983; Gray, 1929; Gueron and Liron, 1993; Taylor, 1951). Flagella, which are similar to cilia but are typically used for cellular propulsion, have also been investigated in tandem. First investigated by Gray (1929) and later in more depth by Taylor (1951), the movement of cilia and flagella in viscous fluid has been modelled with increasing realism. Realistic, 3D computational models of cilia and flagella beating based on the underlying biomechanical structure have been created (Dillon et al., 2007; Gueron and Levit-Gurevich, 2001; Gueron and Liron, 1993). Dillon and Fauci (2000) modelled a single cilium, focusing on the interactions within the internal structure of the cilia. The beating is an emergent property of the internal structure combined with the external fluid dynamics. This work was later extended to include



more realism in fluids with different viscosity for both cilia and flagella, and used to simulate the motion of sperm (Dillon et al., 2007).

Gueron and Levit-Gurevich (2001) created a model of multiciliary beating in 3D based on equations describing the motion, hydrodynamics, internal bending of cilia and the molecular 9+2 microtubule structure of cilia. They observed that the combination of basic behaviours results in coordinated beating of multiple cilia, due to the hydrodynamic coupling between individuals. Multiciliary interactions within fluid were modelled by Yang et al. (2008). An integrative model based on the biomechanical function of the cilia in 2D was created, and showed how fluid viscosity influences the behaviour of cilia. In similar work, Daughton et al. (2008) also investigated the impact of multiciliary interactions on fluid movement, however the ciliary movement is based on digitised observations of real cilia beating, and the emphasis of the study is on the movement of fluid in 3D. These studies helped to show not only how the combined beating of multiple cilia could influence the fluid around them, but also how changes in the fluid can influence the way in which the cilia beat.

Models of sperm movement have been created to investigate behaviour outside the oviduct environment. Hyakutake et al. (2009) created a computational model of a microfluidic sperm sorter. The fluid flow through the sorter is recreated and the distribution of motile spermatozoa between different outlet tubes is investigated based on the flow rate, the swimming patterns and speed of motile sperm, the width of the tubes and the structure of the device. Out of all investigated factors, fluid velocity was found to be a highly significant factor in sperm sorting efficiency.

A curious phenomenon which can be observed in sperm *in vitro* is surface accumulation. This is where sperm will swim close to a solid surface, and even orient themselves to be swimming parallel to that surface. Smith et al. (2009) created a computational simulation to investigate this phenomenon in mammals using accurate and complex theories of hydrodynamics. Different swimming patterns were investigated and sperm swimming with planar beat patterns were found to establish an ‘equilibrium trajectory’ at a specific distance parallel to the surface, resulting in surface accumulation.

Friedrich et al. (2010) used a high-precision sperm tracking method to investigate the validity of early theories on sperm movement against a surface based on resistive force theory (Gray and Hancock, 1955). This theory proposes that the relative frictional forces between sideways and forward movement caused forward propulsion. The sperm head and flagellum were filmed at 250 frames per second (FPS), which allowed a set of accurate data points for the sperm flagellum movements over time to be recorded. When compared with models of sperm movement based on the resistive force theory, the

original theory was shown to have a good match to the accurate measurements. Since resistive force theory was initially proposed, significantly more complex theories have been developed, however this study showed that realistic behaviour can be reproduced using a simpler approximation (Friedrich et al., 2010).

The crawling behaviour of Nematode sperm has also been modelled at the mechanochemical level (Bottino et al., 2002). This finite element model incorporated formulae based on assumptions relating to the underlying biological mechanisms and observations of behaviour. By modelling the low level mechanisms, the resulting behaviour was a good match for the real movement of nematode sperm, providing a basis for future investigation of amoeboid cell motility.

Although the sperm guidance mechanisms in mammals are not fully understood, in amphibians, chemotaxis is a well-established theory. Chemotaxis is the phenomenon in which cells direct their movements according to certain chemicals in their environment. Ishikawa et al. (2004) created a computational model of sperm movement in non-mammals under the influence of chemotaxis. The model simulated sperm from two different species, each with different swimming patterns and different behavioural responses in response to the chemoattractant. Friedrich and Jülicher (2007) created a mathematical description of amphibian sperm swimming patterns in response to chemoattractants.

A wide variety of techniques have been used to model different aspects of the reproductive system, many of which focus on the behaviour of an individual. Clearly, none of the current techniques for modelling sperm behaviour are adequate for investigating the population level behaviour of sperm within the oviduct.

### 3.5 Conclusions

This chapter has provided an overview of computational biology and computational modelling of biological systems. The different techniques and approaches commonly used to create computational models of biological systems and the different areas within reproductive biology which have been modelled computationally have been described.

When looking at the behaviour of gametes in the female reproductive tract, the behaviour of individuals is important. The interaction between the individuals within the system and their environment leads to the regulated passage of sperm to the site of fertilisation. A mathematical model could potentially be created, based on the sperm counts available in literature, to predict the percentage of sperm to reach the site of fertilisation over time. However, this would not allow the inclusion of variations in the

oviductal environment. As the physical 3D structure of the oviduct is an important factor within the system, then a modelling technique that allows this to be included is required.

As described in Section 3.4.1, agent based modelling techniques are well suited for modelling the behaviour of individual cells within a 3D environment. This technique will give the flexibility to include and easily change individual specific behaviour. It will also give the flexibility to change and vary the 3D environment that the individuals move around in. Therefore, the agent based modelling approach will be adopted for the development of the model of sperm behaviour.

Chapter 4 provides a conceptual model of sperm behaviour in the oviduct and Chapter 6 describes the computational implementation of the model.

## Chapter 4

# Defining a model of sperm behaviour

This chapter describes the process of defining the model of sperm behaviour. The modelling process is described and then the domain model, which is the biological description of the system being modelled, is described. The agent based model, which contains the individual processes and parameters of all agents, is then described. Finally, values for the parameters within the model are identified from literature, and the limitations of the available data are discussed. This Chapter provides a link between the literature based description of the biological system, as presented in Chapter 2, and the computational implementation of the model presented in Chapter 6.

### 4.1 Modelling process

Before making a computational model, it is important to establish the purpose of the model. By establishing the purpose, the model can be created to focus on the important behavioural characteristics of the system without introducing unnecessary complexity. It is also important not to try and model everything at once, as the number of variable elements makes it difficult to understand how the different combinations of input values influence the system (Grimm, 1994).

When deciding what to model, it can be useful to investigate something which would be difficult to examine experimentally, possibly due to ethical, financial or practical limitations. With this aim, a good starting point is to initially model aspects of the system which can be validated experimentally. Once validated, the model can then be extended to investigate more complex behaviour.

Ideally, a recursive approach is taken when modelling a biological system. For a given real world problem, a conceptual model is developed. A computational realisation of the conceptual model is then created and tested to ensure that it is an accurate representation of the conceptual model. The computational model can then be experimented

with to validate it against the real world problem, and new real-world experiments can be performed. The results of this experimentation can be used to update the conceptual model, and then the computational model. This process is repeated until a representation of the real world system which is sufficiently accurate to investigate a given problem is established (Ideker et al., 2001; Sargent, 2005). Although ideal, this approach is not always feasible, due to lack of resources and availability of expertise from different fields.

As complementary experimentation was not possible for the behaviour model of sperm due to limited resources, the model is based primarily on qualitative and quantitative data from literature. For the creation of the 3D models of the oviduct, histology images of the oviduct were obtained and the 3D model is based those images. This is discussed further in Chapter 5.

Instead of using complementary experimentation, the model was developed recursively with discussions from experts in the field. An initial prototype of the model was developed and presented to domain experts, along with the limitations of the model. The model was then progressively refined to include more realistic behaviour, based on feedback from domain experts, backed up by data from literature.

## 4.2 Domain Model

This section describes the domain model, which identifies the features of the biological system that will be included. As described in Chapter 2, the oviduct is the part of the female reproductive system which connects the uterus to the ovary and is the site of fertilisation in mammals. The main components of the oviduct are show in Figure 4.1, along with the location of the sperm reservoir in the isthmus and the site of fertilisation in the ampulla.

The domain model presented here represents the final model after the iterative development process was completed. The first version of the model started with only the very basic behaviour. The model was then progressively refined, increasing the complexity and including more mechanisms.

### 4.2.1 Defining the system boundaries

The first step in defining a domain model is to define the system boundaries, that is, the parts of the system which are included in the model. This involves identifying the physical aspects of the system to be modelled, such as the physical environment, individuals within the system and the scale at which they interact, and the tempo-

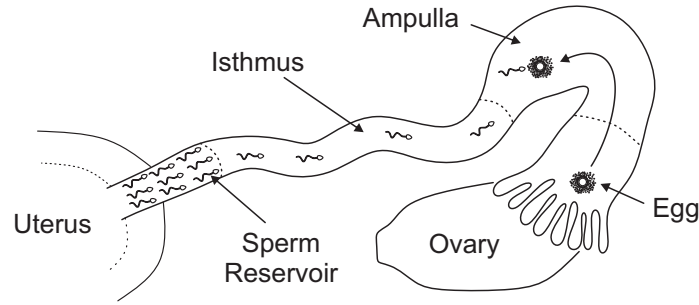


Figure 4.1: The main regions of the oviduct along with an overview of the progression of sperm and oocytes within the system.

ral aspects of the system, such as the time scale that individual interactions will be modelled at and the time period over which the model simulation will cover.

#### 4.2.1.1 Physical environment

As the main aim is to investigate how spermatozoa move within the oviduct, the environment was limited to the section of the oviduct starting just above the UTJ, and finishing at the far end of the ampulla, before the fimbriae.

#### 4.2.1.2 Individuals within the environment

The oviduct contains two main types of individual, spermatozoa and oocytes.

Spermatozoa are autonomous individuals which move on their own using flagellar based propulsion. The behaviour of individual sperm and their interactions with the oviductal environment regulates the passage of spermatozoa from the sperm reservoir to the site of fertilisation (Suarez, 2008b).

The oocytes, which are the female reproductive cells, do not have autonomous movement. After ovulation, oocytes move from the ovary towards the ampulla, primarily due to transport from beating cilia (Lyons et al., 2006). At the time when interactions with sperm are possible they are relatively static at the site of fertilisation (Zamboni, 1972).

#### 4.2.1.3 Temporal environment

As the focus of investigation is the progression of sperm within the oviduct, the model will focus on this region only, and only during the period where spermatozoa and oocyte are capable of meeting, which is the time frame shortly before and after ovulation. The

main goal is for spermatozoa to reach the oocyte. The process of fertilisation once spermatozoa have encountered the oocyte is highly complex and beyond the scope of the model.

### 4.2.2 Defining important behaviour

Chapter 2 provides a description of the different theories surrounding sperm movement and navigation in the female reproductive tract. With so many complex mechanisms potentially influencing the system, including all described mechanisms within a model would lead to too many variables and could make the model difficult to understand or validate. Instead, the main aim was to create a model with a subset of these mechanisms, which can be extended upon in the future. The important aspects of the system, those which have a large amount of supporting evidence, were included.

#### 4.2.2.1 Capacitation

Sperm cells can have two different capacitation states, uncapacitated and capacitated. Uncapacitated spermatozoa are unable to penetrate the oocyte and complete the process of fertilisation (Bedford, 1983; Jaiswal and Eisenbach, 2002). It has been shown that spermatozoa do not become capacitated until they reach the oviduct (Jaiswal and Eisenbach, 2002).

Uncapacitated sperm are stored in a ‘sperm reservoir’, which is just inside the UTJ. This helps to prevent polyspermy by regulating the number of spermatozoa which are actively moving at any one time. As capacitation is required for release from the sperm reservoir (Hunter, 2008), and the sperm reservoir effects the gradual release of a few sperm at a time (Suarez and Pacey, 2006), it is reasonable to assume that spermatozoa within the reservoir will gradually become capacitated over time. In humans, the combination of sperm capacitation and capacitated sperm death was found to result in ‘continuous replacement’ of capacitated sperm (Cohen-Dayag et al., 1995).

While uncapacitated, spermatozoa can live for an extended period of time (Suarez, 2008b). However, once capacitated, spermatozoa only live for a short period of time before they die (Cohen-Dayag et al., 1995; Jaiswal and Eisenbach, 2002).

#### 4.2.2.2 Sperm movement

The most basic mechanism to be included is the movement of spermatozoa and associated changes in swimming patterns, without which the spermatozoa would be unable to meet the oocyte.

As it has been shown that the UTJ prevents or significantly reduces the passage of weak or damaged spermatozoa into the oviduct (Dukelow and Riegler, 1972), it can be assumed that the majority of spermatozoa in the oviduct are motile and capable of reaching the site of fertilisation. Spermatozoa from different species demonstrate a wide variety of swimming patterns and swim at different speeds (Phillips, 1972). However, the majority of swimming patterns from motile sperm can be broadly grouped into two categories, progressive and non-progressive (Goodson et al., 2011). Progressive spermatozoa have a synchronous, regular flagellar beat pattern, which results in straight or slightly curved forward motion. The movement of non-progressive sperm is more difficult to characterise. Non-progressive sperm have an asynchronous, irregular flagellar beat pattern, which can result in a general forward or curved motion with wide path deviation, or a more erratic movement with random changes in direction (Goodson et al., 2011; Ho and Suarez, 2001; Phillips, 1972; Suarez et al., 1991). Capacitated spermatozoa have been observed to spontaneously switch between progressive and non-progressive movement (Suarez, 2008a).

#### 4.2.2.3 Sperm interactions with the oviduct

The wall of the oviduct provides a physical barrier which sperm cells cannot pass through. As sperm cells move forward, they collide with the environment which interrupts their movement pattern. When a collision occurs, the spermatozoon may either attach to the wall, or change direction and swim off (Demott and Suarez, 1992).

Once capacitated spermatozoa are attached to the oviduct, they may detach and swim off in a different direction. This has been observed to primarily occur when they switch to non-progressive movement (Demott and Suarez, 1992).

#### 4.2.2.4 Sperm interactions with oocytes

Sperm swimming in the ampulla region will attach to an oocyte if they collide. The sperm will then start to penetrate the oocyte. In reality the process of fertilisation is far more complex (Ikawa et al., 2010). However, as fertilisation is not the focus of this model, this simple approximation is sufficient.

Spermatozoa will only interact with oocytes which have moved to the site of fertilisation and undergone a maturation process, which could take several hours. Prior to this time, sperm will not attach to the oocyte, and will pass through the ampulla (Braden and Austin, 1954a; Zamboni, 1972).

It has been observed that sperm which meet the mature oocyte will start to penetrate



the Zona Pellucida soon after (Jin et al., 2011). Once the Zona Pellucida has been penetrated, a block to Polyspermy is established (Wolf, 1978). The block is not immediate and multiple sperm can penetrate an oocyte before the block is established.

### 4.3 Defining an agent model of sperm behaviour

It is clear that when modelling sperm populations, the properties and behaviour of individuals are important. The chosen approach was to model sperm at a cellular level using agent based modelling. The individual behaviour of sperm was characterised as a set of rules and the model was used to investigate how the combination of individual behaviours within the oviductal environment results in successful sperm - oocyte encounters.

By implementing a set of rules that encapsulate the mechanisms involved in sperm navigation, we can investigate the relative significance of each mechanism and how the population behaves in relation to the complexity of their environment. Therefore, the model will allow us to investigate how variations in the different factors influencing sperm movement within the oviductal environment at the time of fertilisation affect the distribution of spermatozoa within the oviduct, the time taken for spermatozoa to meet the oocytes and incidence of polysperm.

A generic model of sperm behaviour was created, and species specific variations in behaviour were not included, but could be simulated later by varying the parameters. A subset of the behaviours which are reported in many species are included, with scope to include more complex behaviour later.

This section provides a specification for the agent model. The processes of the model are wrapped in square brackets and shown in [SMALL CAPS], the properties of the model are wrapped in square brackets and highlighted using a [Teletype Font] and the values for properties are marked in *italics*.

Figure 4.2 provides a basic overview of how the processes used within the model are linked together and Figure 4.3 shows the different states that sperm agents can hold and how transitions between those states occur. Table 4.1 shows the system level parameters used in the model that influence sperm and oocyte behaviour and the associated units of measure. System level parameters have a single value for all agents. Table 4.2 shows the agent level parameters used for both spermatozoa and oocytes, and the associated units of measure. Agent level parameters have different values for each agent. Table 4.3 shows the inputs to the model. All assumptions used in the model are listed in Appendix D.

The model created was largely generic, but for the purposes of parameterising the model, the mouse was used as a target species. Mice are commonly used in biological experiments, so there is a large amount of literature published using the mouse. As one of the key aspects of the system is sperm movement within the environment, the model needs two main components, (1) the behaviour model for sperm and (2) a set of 3D models for the oviduct environment itself.

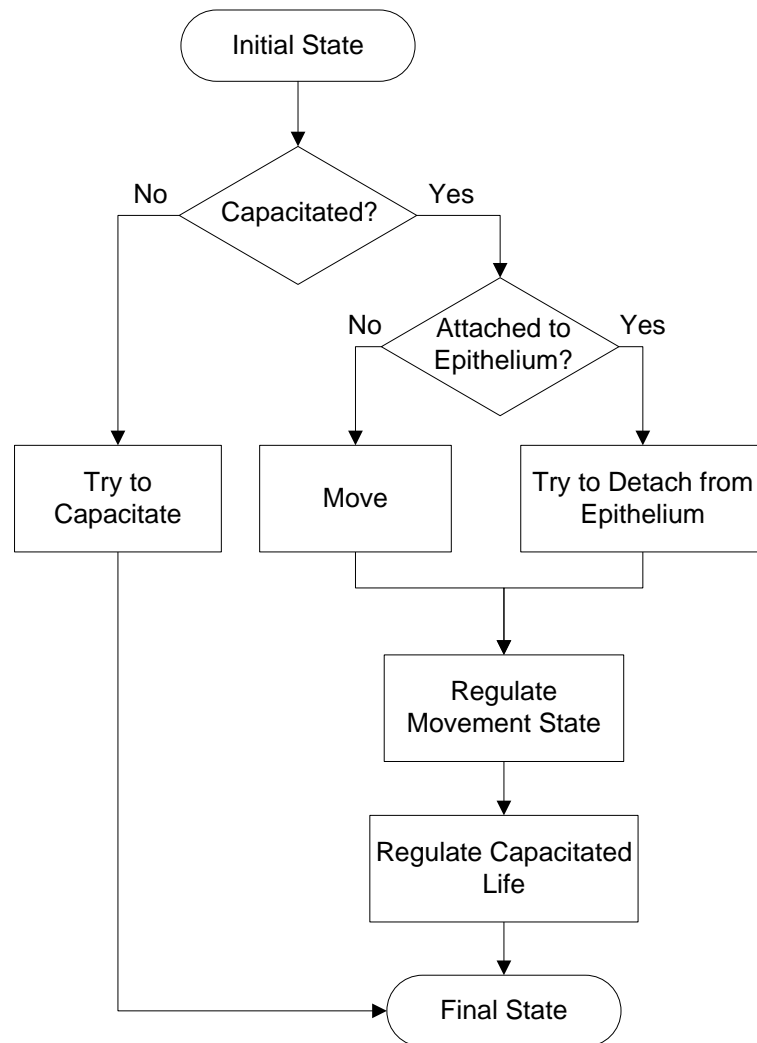


Figure 4.2: The processes followed by each sperm agent every iteration.

#### 4.3.1 Creating the system boundaries

This section describes how the boundaries of the system are implemented.

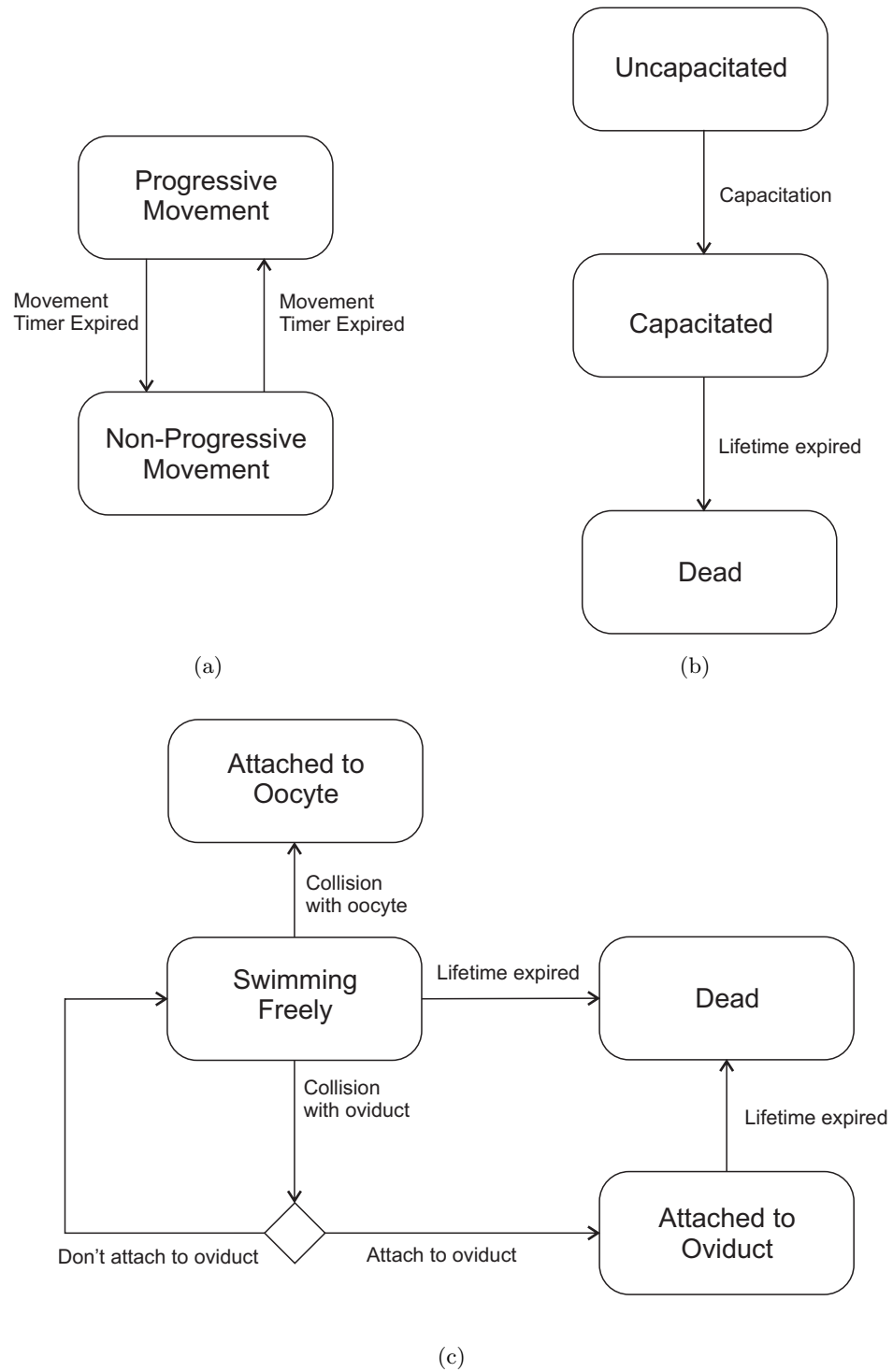


Figure 4.3: The different states individual sperm agents can have and transitions between those states. (a) Sperm can switch between progressive and non-progressive movement. (b) Initially uncapacitated, sperm gradually become capacitated then die once their lifetime has completed. (c) Sperm swimming freely can either attach to the oviduct, attach to an oocyte or die.

<b>Sperm Specific System Parameter</b>	<b>Units</b>
Sperm Radius	Size ( $\mu m$ )
Capacitation Threshold	Probability (P)
Progressive Attachment Threshold	Probability (P)
Non-Progressive Attachment Threshold	Probability (P)
Progressive Detachment Threshold	Probability (P)
Non-Progressive Detachment Threshold	Probability (P)
Time in Progressive State	Time Range (s)
Time in Non-Progressive State	Time Range (s)
Capacitated Sperm Life	Time (s)
Progressive Velocity	Velocity ( $\mu m s^{-1}$ )
Non-Progressive Velocity	Velocity ( $\mu m s^{-1}$ )
No. of Movement Steps	Count (n)
Max. Non-Progressive Angle	Angle (deg)
Max. Reflection Angle	Angle (deg)
Max. Detachment Angle	Angle (deg)

<b>Oocyte Specific System Parameter</b>	<b>Units</b>
Oocyte Radius	Size ( $\mu m$ )
Oocyte Maturation Time	Time (s)
Block to Polyspermy Time	Time (s)

Table 4.1: Sperm and oocyte specific system level parameters used in the model.

<b>Sperm Parameter</b>	<b>Units</b>
Current Position	3D Position [x, y, z] ( $\mu m$ )
Current Direction	3D Normal [x, y, z] (Unit Vector)
Capacitation State	( <i>Uncapacitated</i> or <i>Capacitated</i> )
Alive State	( <i>Alive</i> or <i>Dead</i> )
Movement State	( <i>Progressive</i> or <i>Non-Progressive</i> )
Attachment State	( <i>Attached to Epithelium</i> , <i>Touching Epithelium</i> , <i>Free Swimming</i> or <i>Attached to Oocyte</i> )
Remaining Life	Time (s)
Movement State Timer	Time (s)

<b>Oocyte Parameter</b>	<b>Units</b>
Current Position	3D Position [x, y, z] ( $\mu m$ )

Table 4.2: Agent level parameters used in the model.

Model Input	Units
No. of Sperm	Agent Data (Initial position & direction)
No. of Oocytes	Agent Data (Initial position)
3D Environment	3D Model Data
Total Simulation Time	Time ( <i>s</i> )
Ovulation Time	Time ( <i>s</i> )

*Table 4.3:* Inputs to the model.

#### 4.3.1.1 Physical environment

For investigating the significance of the 3D environment on the behaviour of sperm, a reasonable representation of the 3D environment is required.

In order to represent the environmental bounds described in Section 4.2.1.1, a set of models of the physical 3D environment were created, as described in Chapter 5. The movement of spermatozoa and oocytes is restricted to within the boundaries of the 3D environment.

The 3D environment used in a particular simulation run is an input to the model. The complexity of the environment will be varied to investigate the effects on sperm distribution and oocyte penetration.

#### 4.3.1.2 Individuals within the environment

As described in Section 4.2.1.2, the oviduct contains two main types of individual, spermatozoa and oocytes, both of which are modelled as individual cells.

Spermatozoon are autonomous individuals, which move within and interact with their environment. In order to track this movement, all sperm have agent level parameters representing their [Current Position] in the 3D environment and their [Current Direction]. To simplify the interactions, only the sperm head is modelled, using a sphere. The radius of the sphere representing the spermatozoon ([Sperm Radius]) is a system level parameter, the value of which can be derived from measurement of the width of a sperm head. The oviduct 3D model acts as a physical barrier only, and interactions between individual epithelial cells and sperm cells are not explicitly modelled. The initial position of sperm within the oviduct is in the ‘sperm reservoir’, which is in caudal isthmus, as shown in Figure 4.1.

The oocytes are represented by spheres, which have a fixed position in space at the

site of fertilisation. Although in reality oocytes do move down the oviduct, at the time when interactions with sperm are possible they are relatively static at the site of fertilisation (Zamboni, 1972). In order to identify their position, oocytes have an agent level parameter representing their `[Current Position]` in the 3D environment. The radius of the sphere representing an oocyte (`[Oocyte Radius]`) is a system level parameter, the value of which can be derived from measurements of the width of an oocyte without the OCC. Interactions between oocytes and spermatozoa are modelled at a cellular level. The initial position of the oocytes within the oviduct is at the site of fertilisation in the ampulla, as shown in Figure 4.1.

Spermatozoa and oocytes both have a limited maximum lifetime, but it is assumed that the simulation will end before this lifetime is reached (Harper, 1982).

#### 4.3.1.3 Temporal environment

The total amount of time the simulation should run for, called the `[Total Simulation Time]`, is an input parameter. The value of this parameter can be determined from studies which look at the amount of time taken for spermatozoa to reach or fertilise oocytes with respect to time, relative to ovulation.

The time at which ovulation occurs relative to the start of the simulation (`[Ovulation Time]`) is an input parameter.

#### 4.3.1.4 Timing of interactions

Spermatozoa move very fast in relation to their size, and their flagellar beat frequency is several beats per second (Goodson et al., 2011). Modelling at an individual beat level is possible, but this would result in a large number of calculations, generate large amount of data, and significantly increase the time required to complete a single simulation run without any significant benefit. Also, many of the parameters and measurements reported in literature are relative to one second intervals. Due to these factors, the simulation time step was set to one second per iteration, and all parameters used within the system are relative to this interval.

### 4.3.2 Agent behaviour

This section describes the processes related to individuals within the system, along with system parameters, agent level parameters and the inputs to the model.

As identified in Section 4.2.2, the individual behaviours included in the model are sperm

capacitation, capacitated sperm death, sperm movement, sperm interactions with the oviduct, detachment from the epithelium and sperm interactions with oocytes.

#### 4.3.2.1 Capacitation

To represent the sperm reservoir within the model, all spermatozoa start in an uncapacitated state.

Sperm capacitation is a process of the model (`[TRY TO CAPACITATE SPERM]`) that determines if currently uncapacitated sperm should become capacitated at the current iteration. The process is encoded as a probability (`[Capacitation Threshold]`), the value of which was estimated based on the assumption that the majority of uncapacitated sperm will have capacitated within a specific time period.

The probability ( $P$ ) that a single uncapacitated spermatozoon will become capacitated at any second is calculated in relation to the number of hours ( $ct$ ) that the majority should be capacitated by as:

$$P = \frac{1}{(ct \times 3600)} \quad (4.1)$$

with 3,600 being the number of seconds in an hour. The number of hours for all sperm to become capacitated ( $ct$ ) can be estimated from the total sperm lifetime for a particular species. The `[Capacitation Threshold]` is a system level parameter, and the sperm `[Capacitation State]` is an agent level parameter. The different capacitation states are shown in Figure 4.3(b).

#### 4.3.2.2 Capacitated sperm death

All sperm cells are initially *alive* (`[Alive State]`), but once capacitated, spermatozoa only live for a short period of time before they die (Cohen-Dayag et al., 1995; Jaiswal and Eisenbach, 2002). This is different from the sperm total lifetime, which is related to the maximum life of sperm measured from initial deposition until the last sperm stops moving.

Sperm death is a process of the model (`[REGULATE CAPACITATED SPERM LIFE]`) that determines if currently capacitated sperm should die. When an uncapacitated sperm becomes capacitated, the length of time that it can live for in seconds (`[Remaining Life]`) is set. This is reduced by one every iteration until it reaches zero, at which point the `[Alive State]` is set to *dead*.

The amount of time a spermatozoon can live for ([Capacitated Sperm Life]) is a system level parameter and the [Remaining Life] and [Alive State] are agent level parameters. The different capacitation states are shown in Figure 4.3(b).

#### 4.3.2.3 Sperm movement

Section 4.3.2.3 describes the different movement states that spermatozoa present. All spermatozoa have an agent level parameter representing this [Movement State], which can hold the values *progressive* and *non-progressive*. If spermatozoa are in the *progressive* movement state, then they follow the [PROGRESSIVE MOVEMENT] process. Within this process, the [Current Position] is updated to reflect the agents moving forward in their [Current Direction] at a fixed rate, determined by the [Progressive Velocity] system level parameter. If spermatozoa are in the *non-progressive* movement state, then they follow the [NON-PROGRESSIVE MOVEMENT] process. In order to simulate non-progressive movement, sperm agents randomly change direction and then move forward.

The magnitude of rotation is randomly calculated to be between zero and [Max.-Non-Progressive Angle], which represents the maximum number of degrees the sperm cell can change direction every second, and applied to the [Current Direction]. The rotation is conic, as illustrated in Figures 4.4(a) and 4.4(b), where ( $\theta$ ) is the randomly determined rotation angle. The [Current Position] is then updated to reflect the agents moving forward in their [Current Direction] at a fixed rate, determined by the [Non-Progressive Velocity] system level parameter.

Switching between progressive and non-progressive movement states is handled in the [REGULATE MOVEMENT STATE] process. Uniform distributions encapsulating the amount of time sperm spend in each movement state were identified ([Time in-Progressive State] & [Time in Non-Progressive State]). When sperm switch movement state, then the amount of time they will spend in that state is randomly selected from the corresponding uniform distribution, and stored in the [Movement-State Timer] agent level parameter. Every iteration, the value of this parameter is reduced by one. When it reaches zero, the movement state is switched, and a new [Movement State Timer] value is calculated. The different movement states are shown in Figure 4.3(a).

#### 4.3.2.4 Sperm interactions with the oviduct

Section 4.3.2.4 describes the interactions between spermatozoa and the oviductal environment. All spermatozoa have an agent level parameter called the [Attachment



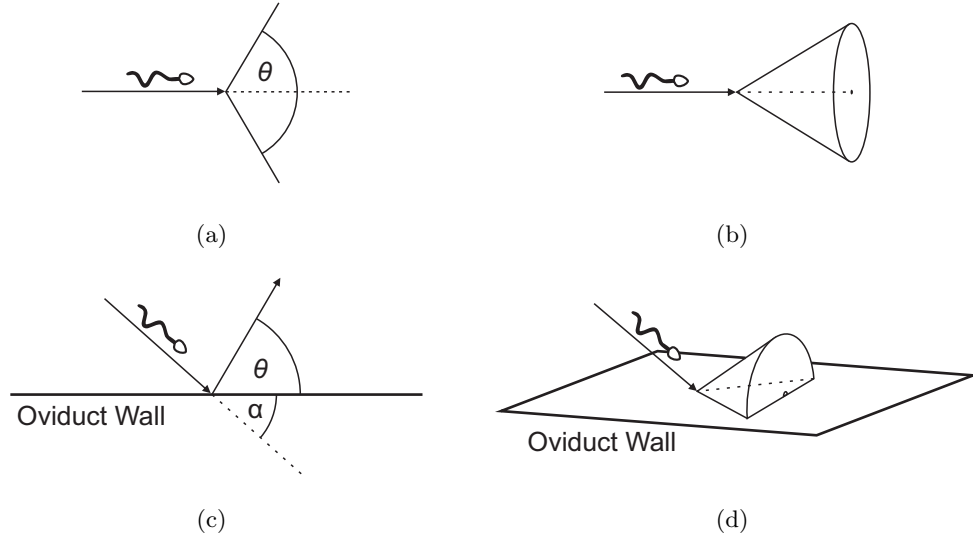


Figure 4.4: Differences between the rotation and reflection methods used within the model. (a) Non-progressive sperm rotate randomly between 0 and  $\theta$  and (b) the resulting potential rotation cone in 3D. (c) Sperm interacting with the oviduct rotate  $\alpha$  degrees to be parallel to the surface, then randomly between 0 and  $\theta$  and (d) the resulting potential rotation half-cone in 3D.

**State]**, which represents the possible states of interaction with the environment. The **[Attachment State]** can hold the values *attached to the oviduct*, *touching but not attached*, *free swimming in the lumen* or *attached to oocyte*. Interactions between sperm and oocytes are described in Section 4.3.2.5.

Within the movement process, collisions between moving sperm and the oviduct wall are detected, and the **[COLLISION RESOLUTION]** process is followed. This process determines if a sperm which collides with the oviduct should attach or reflect off in another direction. This is encoded as a probability (**[Progressive Attachment Threshold]**), the value of which is based on observational data about how frequently sperm attach to the wall when contact is made. The **[Progressive Attachment Threshold]** will have different values, depending on whether the **[Movement State]** property has the value of *progressive* or *non-progressive*. If an attachment is made, then the **[Attachment State]** is set to *attached to epithelium*.

If an attachment is not made, then the **[Attachment State]** is set to *touching epithelium*, and the sperm rotates to reflect off in a new direction. The sperm initially rotates until it is parallel with the oviduct wall. A second rotation is then applied, the magnitude of which is randomly calculated to be between zero and **[Max. Reflection Angle]**, which represents the maximum angle the sperm can reflect off the oviduct surface,

and applied to the **[Current Direction]**. The rotation is half-conic, as illustrated in Figures 4.4(c) and 4.4(d), where  $(\alpha)$  is the angle initially rotated until the sperm is parallel with the oviduct wall and  $(\theta)$  is the calculated angle of rotation.

As sperm move quickly in relation to their size every second, during a single iteration, they could potentially have multiple collisions with the environment. To represent this, the movement is subdivided into smaller steps per iteration, which allows for multiple collisions. The number of movement steps per iteration (**[No. of Movement Steps]**) is a system level parameter, the value of which can be determined by estimating the number of collisions a single sperm can make each second. At the end of the movement process, if sperm are not attached to or touching the oviduct, then the **[Attachment State]** is set to *free swimming*.

Sperm that have an **[Attachment State]** of *attached to the oviduct* will stay attached and not move until they detach. The **[TRY TO DETACH FROM EPITHELIUM]** process determines whether or not sperm should detach from their oviductal bindings and continue to swim freely within the oviduct. This process is encoded as a probability, the value of which can be determined from observations on how frequently attached sperm detach from the oviduct.

The probability will have different values for progressive (**[Progressive Detachment Threshold]**) and non-progressive (**[Non-Progressive Detachment Threshold]**) movement, depending on whether the **[Movement State]** property has the value of *progressive* or *non-progressive*.

If a detachment is made, then the **[Attachment State]** is set to *touching epithelium*, and the sperm rotates to point in a new direction. The magnitude of rotation is randomly calculated to be between zero and **[Max. Detachment Angle]**, which represents the maximum angle the sperm can reflect off the oviduct surface, and applied to the **[Current Direction]**.

The way a new direction is calculated is similar to when sperm reflect off the oviduct wall and do not attach, the key difference being the maximum angle of rotation. This is shown diagrammatically in Figures 4.4(c) and 4.4(d). If a detachment is not made, then the sperm remain attached and do not change direction. The **[Progressive Detachment Threshold]** and **[Non-Progressive Detachment Threshold]** are a system level parameters. The different sperm interaction states are shown in Figure 4.3(c).

#### 4.3.2.5 Sperm interactions with oocytes

Section ?? describes the interactions between spermatozoa and oocytes. Within the movement process, collisions between moving spermatozoa and the oocytes are detected, and the [Oocyte Collision] process is followed.

The ([Oocyte Maturation Period]) is a system level parameter that represents the amount of time required after ovulation before sperm are able to penetrate the oocyte. Prior to this, a spermatozoon which collides with an oocyte will pass through without collision. When a spermatozoon does collide with an oocyte, its [Current Position] is set to the collision point and its [Attachment State] is set to *attached to oocyte*. Once in this state, the sperm agent will remain attached to the oocyte for the remainder of the simulation. The different sperm interaction states are shown in Figure 4.3(c).

#### 4.3.2.6 Sperm to sperm interactions

Agent based models commonly simulate the interactions between individuals within the system. However, for this system, the interactions between sperm agents and their environment are the most significant factor. Sperm to sperm interactions are not modelled, and sperm effectively pass through each other. In the wood mouse, sperm have been shown to attach together at the head, forming long trains (Moore et al., 2002). However, these long trains begin to disperse once they move from the vagina to the uterus, and are completely dispersed once capacitation occurs (Moore et al., 2002). Due to the small number of sperm which are capacitated at any one time, and the large size of the oviductal environment in relation to the sperm agents, interactions between individual sperm are unlikely to have a significant impact on the progression of sperm.

#### 4.3.3 Inputs and outputs

The inputs to the model (initial conditions and simulation length) are the sperm agents with their initial position, direction and state; the oocytes agents with their position in the environment; the total simulation time and the system ovulation time. The system outputs (measured emergent behaviour) are the position and state of sperm at different time points; the number of oocytes with sperm attached and the number of sperm attached to each oocyte.

Although fertilisation is not explicitly modelled, incidence of polyspermy can be calculated. Polyspermy occurs when multiple sperm penetrate a single oocyte. It has been observed that it only takes a short period of time for an attached sperm to start to penetrate the Zona Pellucida (Austin and Braden, 1956; Jin et al., 2011) and that the

first sperm to reach the oocyte is likely to be the one that fertilises it (Zamboni, 1972). Therefore, it is assumed that the first sperm to reach the oocyte also penetrates it. Once penetrated, a block to polyspermy is established, however this process is not instantaneous in mammals (Wolf, 1978). The polyspermic block prevents penetration but it does not prevent attachment (Maluchnik and Borsuk, 1994). Therefore, it is assumed that sperm reaching an oocyte which has already established the block to polyspermy will still attach. Incidence of polyspermy can then be calculated based on the average time taken for a block to polyspermy to be established ([Block to Polyspermy Time]), and the number of sperm attaching to the oocyte within that time period. The amount of time taken for the block to polyspermy to be established is available from literature. Therefore, the number of sperm entering a single oocyte can be estimated as the number of attachments within the period between the first attachment and the time of the block.

From these output measures the following results can be calculated:

- The number of penetrated oocytes at specific time points.
- The number of sperm in each oviductal region at specific time points.
- The number of polyspermic oocytes at specific time points.
- The number of capacitated sperm in the oviduct at specific time points.

#### 4.4 Populating the model with parameters

Once a generic conceptual model has been defined, it needs to be parameterised with real data. The data available from literature can be broadly split into two categories: system level parameters and individual parameters. System level parameters relate to observations and measurements of the system as a whole, whereas individual parameters relate to observations and measurements of individual cells within the system.

Typically, individual observations and measurements were used to parameterise the model, and some of the system level observations and measurements were used to define the initial conditions and system level events. The remainder of the system level observations and measurements were used for validation. Importantly, different sources are used for defining the model and validating the model.

System level observations from literature are usually statements which generalise the behaviour of the system. Many of these observations are qualitative or speculative. Ideally, a computational model requires quantitative data. Where this is not available, estimates of the parameter values are made based on qualitative descriptions. Where

the value for a particular parameter is uncertain, the full range of potential values was identified and an initial value was chosen. A large range of values for each uncertain parameter can then be tested within the full range during the model calibration phase.

In order to make the computational model more realistic, it was tailored to data from a single species. Mice are commonly used for biological experiments, due to their low cost in relation to other animals, a short oestrous cycle, availability and more relaxed ethical constraints in relation to larger mammals and humans. Because of this, a large amount of literature and data are available for mouse reproduction and fertility. Therefore the mouse was chosen as a target species for the computational model.

This section describes the types of data available for the mouse from different sources in literature, how they are used to parameterise the model and their limitations. Table 4.4 shows the initial value and range for each system level parameter used in the model and Table 4.5 shows the initial model inputs used to test the model. The timeline for the first few hours of the reproductive cycle in mice is shown in Figure 4.5.

#### 4.4.1 Initial conditions

This section describes how the initial conditions were determined from literature.

##### 4.4.1.1 Initial simulation time and sperm distribution

After coitus, spermatozoa are transported rapidly to the uterus (Suarez and Pacey, 2006) and the UTJ, which is normally closed, opens for approximately 1 hour to allow passage of spermatozoa into the oviduct (Suarez, 1987; Zamboni, 1972). After this time the UTJ closes, preventing new spermatozoa from entering the oviduct and preventing spermatozoa from leaving the oviduct (Hafez and Black, 1969; Suarez, 1987; Zamboni, 1972). Once inside the oviduct, spermatozoa attach to the epithelium of the intramural and extramural isthmus (Nicol and McLaren, 1974; Suarez, 1987; Zamboni, 1972). In mice, it has been suggested that sperm require one hour in the FRT to become capacitated (Braden and Austin, 1954a; McGaughey et al., 1968).

From this information, we can initialise our model with sperm agents distributed along the inner wall of the caudal isthmus, oriented away from the UTJ. The simulation starting time will be 1 hour *post coitus*, which is the time point when sperm are no longer able to enter the oviduct and the time when sperm can start to become capacitated.

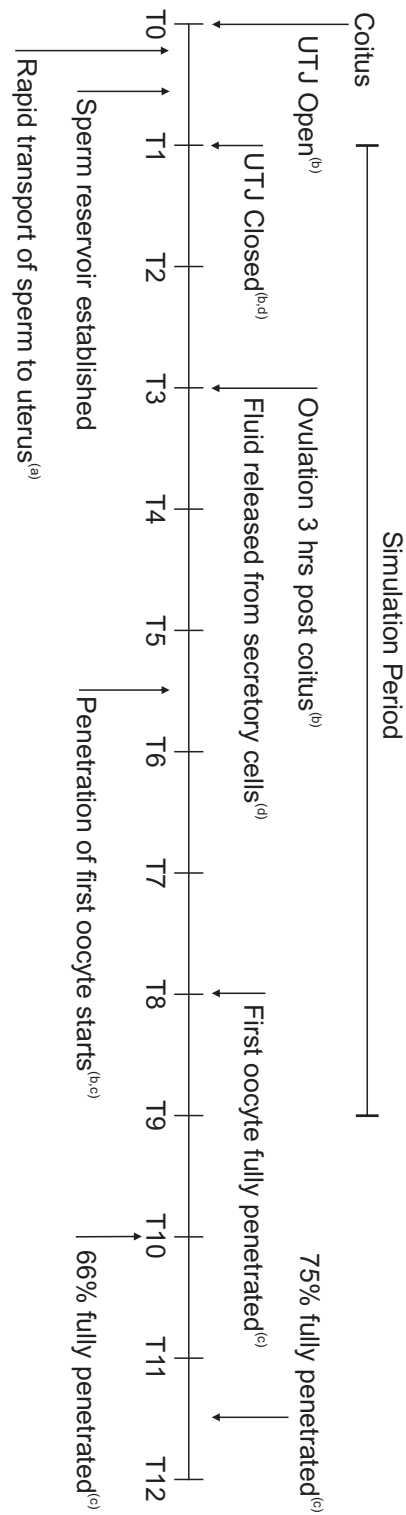


Figure 4.5: The timeline of events at the start of the mouse reproductive cycle. The model represents the period of time from when the UTJ closes 1h *post coitus* and lasts for eight hours. (a) Suarez and Pacey (2006) , (b) Zamboni (1972), (c) Braden and Austin (1954a), (d) Perkins (1972), (e) Harper (1982), (f) McGaughey et al. (1968)

Sperm Specific System Parameter	Value	Source
Sperm Radius	1.6 $\mu m$	Measured (a)
Capacitation Threshold*	4.63E-05	Estimated (c)
Progressive Attachment Threshold*	0.5	Estimated (d)
Non-Progressive Attachment Threshold	0.0	Observed (d)
Progressive Detachment Threshold	0.0	Observed (d)
Non-Progressive Detachment Threshold	1.0	Observed (d)
Time in Progressive State	10s – 60s	Observed (d)
Time in Non-Progressive State	5s – 10s	Observed (d)
Capacitated Sperm Life	7200s	Anecdotal (e)
Progressive Velocity	146.9 $\mu m s^{-1}$	Measured (h)
Non-Progressive Velocity	73.3 $\mu m s^{-1}$	Measured (h)
No. of Movement Steps*	4	Estimated (h)
Max. Reflection Angle*	45°	Estimated (d)
Max. Non-Progressive Angle*	90°	Estimated
Max. Detachment Angle*	90°	Estimated (d)

Oocyte Specific System Parameter	Value	Source
Oocyte Radius	36.65 $\mu m$	Measured (b)
Oocyte Maturation Period	9000s	Observed (f & g)
Block to Polyspermy Time	2400s	Observed (i)

Table 4.4: Initial values of sperm and oocyte specific system level parameters used in the model which were identified from literature for the mouse. The parameters with estimated or uncertain values are marked with an asterisk. (a) Firman and Simmons (2010), (b) Kim et al. (2004), (c) McGaughey et al. (1968), (d) Demott and Suarez (1992), (e) Bedford (1983), (f) Zamboni (1972), (g) Braden and Austin (1954a), (h) Goodson et al. (2011), (i) Wolf (1978)

Parameter	Value	Source
Total Simulation Time	28800s	Estimated (a)
No. of Sperm	1536	Average (b)
No. of Oocyte	4	Average (a)
Ovulation Time	7200s	Anecdotal (c)
Sperm Population	1536	Average (b & d)
3D Environment	OvA1	(e)

Table 4.5: Initial values of inputs used in the model which were identified from literature for the mouse. (a) Braden and Austin (1954a), (b) Hodges et al. (2008), (c) Zamboni (1972), (d) Nicol and McLaren (1974), (e) Generated as described in Chapter 5.

#### 4.4.1.2 Ovulation time

Coitus has been said to occur approximately three hours before ovulation in nature, which is when the females are most receptive to mating (Zamboni, 1972). As the simulation start time is 1 hour *post coitus*, the initial [Ovulation Time] was set to two hours.

#### 4.4.1.3 Number of sperm

The number of spermatozoa within the mouse oviduct has been widely reported, with counts ranging from a few to many thousand (Bennett and Dunn, 1967; Braden, 1958; Braden and Gluecksohn-Waelsch, 1958; Harper, 1982; Hodges et al., 2008; Nicol and McLaren, 1974; Olds, 1970; Tessler and Olds-Clarke, 1981). For some of the older studies, the methods used to isolate the sperm population within the oviduct may have been flawed, as they cut the tube above the uterus (Austin, 1952; Bennett and Dunn, 1967; Braden, 1958; Olds, 1970). As we now know that the majority of the spermatozoa within the sperm reservoir are trapped in the most caudal part of the isthmus which starts within the uterine wall, this technique was likely to exclude a large number of spermatozoa which are within the oviduct, which can also account for the wide variation in sperm count between individuals. More recent studies report larger sperm populations within the oviduct (Hodges et al., 2008), but large variations between individuals and individual studies still exist. The initial [No. of Sperm] was chosen to be 1,536, which was the mean average number of sperm reported in a recent study by Hodges et al. (2008).

#### 4.4.1.4 Number of oocytes

Within the oviduct, the number of oocytes per oviduct is also reported with significant differences. One study which uses a large number of mice reported the average number of oocyte per mouse to be 8.1, implying around 4 per oviduct (Braden and Austin, 1954a). However, other studies report the number of oocytes per oviduct to be far greater (Tessler and Olds-Clarke, 1981). Therefore, the initial [No. of Oocytes] was chosen to be 4.

#### 4.4.1.5 3D environment

The 3D environments were constructed based on measurements taken from histology images. The full process is described in Chapter 5.



#### 4.4.2 System level parameters

This section describes how values for the system level parameters were determined from literature. The term ‘system parameter’ is used to describe parameters where the value is constant throughout the entire system, regardless of whether the parameters is a property of an individual or of the system as a whole.

##### 4.4.2.1 Sperm radius

The size of spermatozoa and oocyte cells are available from multiple sources in literature. Massive comparative studies of sperm cell morphology for different species have been reported and measurements for different components of the spermatozoon are available (Cummins and Woodall, 1985; Firman and Simmons, 2010). As individual spermatozoon are modelled as spheres, their diameter is taken to be the average spermatozoon head width, which set the [Sperm Radius] to  $1.6\mu m$ .

##### 4.4.2.2 Oocyte radius

The diameter of oocytes, which are also modelled as spheres, has been estimated by averaging measurements of the oocyte without the OCC obtained from images found in literature (Kim et al., 2004), with the [Oocyte Radius] set to  $36.65\mu m$ .

##### 4.4.2.3 Capacitation threshold

McGaughey et al. (1968) inseminated mice with sperm at different time points prior to ovulation. They found that spermatozoa inseminated 12-14 hours before ovulation were able to fertilise some oocytes, but the total number of oocytes which were fertilised was significantly reduced once the time between insemination and ovulation became greater than 4-6 hours (McGaughey et al., 1968). This implies that the majority of sperm should be capacitated by 4-6 hours, but that some sperm could still become capacitated after 12-14 hours. As the majority of sperm should be capacitated within 6 hours, the value for the [Capacitation Threshold] is calculated as described in Section 4.3.2, with (ct) having an initial value of 6.

##### 4.4.2.4 Attachment and detachment thresholds

Qualitative observations about the behaviour of individual spermatozoon are also available in literature. Demott and Suarez (1992) reported that upon observing sperm interactions with the oviductal epithelium, “Sometimes the sperm would reattach the first time it contacted the epithelium, and sometimes it would bounce off, frequently mak-

ing a direction change, and then continue swimming freely”. From this, the probability of attaching to the epithelial when spermatozoa collide with the wall (**[Progressive Attachment Threshold]**) was initially estimated to be 0.5, with a full potential range of [0-1]. As this parameter is based on a qualitative description, it was marked as being uncertain.

Mouse spermatozoa have been observed to only detach from the epithelium when non-progressive movement is presented (Demott and Suarez, 1992; Ho et al., 2009). Based on this observation, the probability of detaching from the oviduct when in progressive movement (**[Progressive Detachment Threshold]**) is set to 0 and the probability of detaching when in non-progressive movement (**[Non-Progressive Detachment Threshold]**) is set to 1. As sperm always detach when in non-progressive movement, it was also assumed that sperm would only attach to the oviduct when in the progressive movement state. Therefore, the **[Non-Progressive Attachment Threshold]** was set to 0.

#### 4.4.2.5 Time in progressive and non-progressive state

Demott and Suarez (1992) observed mouse sperm to spontaneously switch between progressive and non-progressive movement. They noted that progressive movement usually occurred for between 10-60 seconds, followed by a burst of non-progressive movement for between 5-10 seconds. Therefore, the **[Time in Progressive Movement]** and **[Time in Non-Progressive Movement]** ranges were set to [10-60] and [5-10] respectively.

#### 4.4.2.6 Capacitated sperm life

Capacitated sperm have been observed to only live for around 2 hours (Bedford, 1983). From this, the **[Capacitated Sperm Life]** was set to 2 hours.

#### 4.4.2.7 Oocyte maturation time

It has been observed that the period between ovulation and the start of penetration of oocytes is between 2.5-3 hours (Braden and Austin, 1954a; Zamboni, 1972), leading to an initial value for the **[Oocyte Maturation Time]** of 2.5 hours.

#### 4.4.2.8 Progressive and non-progressive velocity

The measurable movement characteristics of spermatozoa are commonly reported in literature using Computer Assisted Semen Analysis (CASA) (Firman and Simmons, 2010; Goodson et al., 2011; Grzmil et al., 2007; Mortimer, 2000). This is where spermatozoa

samples are videotaped using high frequency video cameras attached to microscopes. This technique provides a quick snapshot of the behaviour of individual sperm, with typical recordings of individual spermatozoon lasting for only a few seconds. The video footage is analysed, and usually the path that the head follows is identified. From this, a set of threshold values can be defined to categorise the sperm movement patterns. A large number of CASA metrics exist, some of which are calculated from combinations of other metrics. The main characteristics which are measured are the speed, lateral head movement and linearity, which is the curved movement path that a spermatozoon follows over time (Mortimer, 2000). When repeated for hundreds of spermatozoa, the different types of movement patterns and velocities of spermatozoa in different states can be identified. This gives a good overview of the variations in sperm movement, but care needs to be taken when using this as a generalisation for how an individual spermatozoon moves over its entire lifetime.

When calculating the movement speed, three different metrics are commonly used, Straight Line Velocity (VSL), Curvilinear Velocity (VCL) and Average Path Velocity (VAP) (Mortimer, 2000). As spermatozoa move forward with a series of flagellar beats, their head moves back and forth and rotates as the spermatozoa move forward. When the position of the head only is measured, then the path of a progressive spermatozoa appears as a snaking line. The VCL is based on the total head movement, following the curved path. Another method used is the VSL, which is based on the total distance moved over the recorded period, measuring a straight line from the start point to the end point. The VAP is a combination of the two, which is an estimate of the average movement path. As the progressive movement of spermatozoa within the model is based on the distance moved per second, the VAP is used. For non-progressive sperm, the VSL of hyperactivated spermatozoa is used as an approximation for how far they should move.

The accuracy of CASA in characterising the movement of spermatozoa is heavily dependent on numerous factors. Fluid viscosity, temperature, capacitation state, capacitation medium and liquid depth can all influence the measurements obtained. The accuracy of CASA is also dependent on technical factors, such as the image resolution, camera frequency (frames per second) and the threshold parameters used to characterise different motility patterns (Mortimer, 2000). When looking at reported CASA parameters for mouse sperm, significant differences between studies are observed. Firman and Simmons (2010) report the average VAP of progressive spermatozoa to be  $52\mu\text{ms}^{-1}$ . However other studies report the average VAP of progressive spermatozoa to be  $146.9\mu\text{ms}^{-1}$  (Goodson et al., 2011),  $117.9\mu\text{ms}^{-1}$  (Tessler and Olds-Clarke, 1985) and  $135\mu\text{ms}^{-1}$  (estimated from chart) (Suarez and Dai, 1992), which are all

significantly faster. When looked at in detail, the studies use different strains of mice, different capacitation medium, the cells are incubated for different time lengths and different CASA thresholds are used for analysing sperm motility, which can account for the differences in measurements.

The chosen values for **[Progressive Velocity]** and **[Non-Progressive Velocity]** were taken from Goodson et al. (2011), which gave values of  $146.9\mu\text{ms}^{-1}$  and  $73.3\mu\text{ms}^{-1}$ . This study was based on several thousand measurements, so can be considered to be an accurate representation of the sperm movement speed.

#### 4.4.2.9 Number of movement steps

One of the parameters with little information from literature is the number of times a spermatozoon can reflect and continue moving each second (**[No. of Movement-Steps]**). Sperm movement is powered by a series of flagellar beats, with progressive sperm having an average cross beat frequency (CBF) of 25.4 per second (Goodson et al., 2011). The CBF is the number of times the flagellar is observed to cross the centreline of the sperm path, so the number of full beats is half the CBF. Logically, as sperm move a long way in relation to their size every second, they could potentially reflect multiple times off the surface each second. If we assume it should take at least 1 full beat to turn away from a surface once a collision has occurred, then this means each reflection should use two full beats, one to reach the surface and another to turn away. This gives a maximum number of potential reflections per second of 6.35 and a minimum of 1. When rounded up the total range for this parameter is [1-7] and the initial value of 4 was chosen.

As each time a collision is made there is the potential for an attachment to be made, then a large number of collisions is less likely to occur within a single iteration. If this parameter was varied, it is expected that the most significant difference should be between 1 and 2 collisions, with the significance dropping as the number becomes greater.

#### 4.4.2.10 Rotation and reflection angles

The angle of rotation for non-progressive sperm (**[Max. Non-Progressive Angle]**) also has an uncertain value. The potential range for this parameter is [0-180], so the initial value was taken to be 90.

The angle that the sperm swim off in relative to the intersection surface if an attachment was not made (**[Max. Reflection Angle]**) was based on the qualitative description

provided by Demott and Suarez (1992), who reported that “Sometimes the sperm would reattach the first time it contacted the epithelium, and sometimes it would bounce off, frequently making a direction change, and then continue swimming freely”.

The point at which the sperm collides with the oviduct can be identified as a plane, and the reflection angle off the plane has a potential full range of  $[0 - 180]$  degrees. Sperm are very long in relation to their width, and the sperm head does not rotate a large amount during progressive movement. Combined with the narrow swimming space in the lumen of the oviduct, it was assumed that sperm would not be able to turn back on themselves and swim in the direction they originated from on collision with the surface. This is consistent with observations by Nilsson and Reinius (1969), who stated that the narrow lumen may be necessary to help sperm maintain their direction. This assumption limited the maximum reflection angle possible range to  $[0 - 90]$  degrees, and the initial value was estimated to be 45.

Demott and Suarez (1992) noted that on detachment of sperm from the epithelium “The direction in which a sperm swam after release appeared to be random”. As the detachment of sperm is in part due to the erratic movement of sperm during non-progressive movement, then the sperm is more likely to swim off in a random direction. Because of this, the full potential range of the (**[Max. Detachment Angle]**) was identified as  $[0-180]$  degrees, within an initial value identified as 90 degrees.

#### 4.4.2.11 Block to polyspermy

Wolf (1978) described that *in vitro*, a significant drop in the incidence of polyspermy was observed after 40 mins. Therefore, the **[Block to Polyspermy Time]** was set to 2,400 seconds.

#### 4.4.2.12 Total simulation time

Braden and Austin (1954a) observed that the amount of time required for 75% of oocytes to have sperm attached is around 11.5 hours, assuming coitus occurs 3 hours before ovulation. As it has also been observed that it takes around 2.5 hours for a sperm to fully penetrate an oocyte, this means that the amount of time required for 75% of oocytes to have sperm attached can be determined to be 9 hours. As the simulation starts at 1 hour *post coitus*, then total simulation time before 75% of oocytes have sperm attached is 8 hours, or 28,800 seconds.

## 4.5 Conclusions

The main aim of this chapter was to provide a detailed description of the conceptual model of sperm behaviour. Individual mechanisms and parameters have been identified from literature, and the order in which they are applied to agents within the model has been established. Finally, a set of initial values for all parameters have been determined from literature, based on the mouse. Chapter 5 describes the process used to construct a set of 3D models of the mouse oviduct, and Chapter 6 describes the computational implementation of the agent based model.

## Chapter 5

# Creating a 3D virtual oviduct

This chapter describes the techniques developed to construct a set of 3D models of the oviduct for use with the agent based model. As discussed in Chapter 4, the mouse has been identified as a target species for the model. Therefore, the 3D models for use in the simulation were to be based on the mouse. The structure of the mouse oviduct is described in detail and the level of detail required in the 3D model is established. Related work in 3D construction of biological organs is reviewed, and a suitable reconstruction approach is identified. Histology images of the oviduct were processed to identify the external form and features of the internal structure. These data were then used to construct a set of 2D cross sections which were then combined together to follow the path of the oviductal tube using a novel particle physics inspired approach to resolve intersections. A set of 3D models were constructed using these techniques, which are optimised for display and manipulation in 3D. The limitations of the approach and of the generated models are then identified, and the potential of the developed techniques for use on other species is discussed. Finally, optimisation of the image processing and construction approaches using a combination of multi-core CPU and GPU acceleration is presented.

### 5.1 3D reconstruction of biological organs

Creating realistic 3D computational models of complex biological structures can be a time consuming, manually intensive and computationally expensive process. Existing techniques focus on creating an exact reproduction of a biological sample (Li et al., 2008). However for computational modelling of the biological interactions which occur within a 3D environment, it can be more useful to have a technique that results in a set of unique 3D models, which are not exact copies of the original tissue, but are grounded in reality and accurate enough to provide a realistic environment. This is especially useful for biological systems where the structure is highly complex, with large variations

between individuals and between different species.

Medical imaging techniques such as Magnetic Resonance Imaging (MRI), Computed Tomography (CT) and 3D Ultrasound are commonly used to create 3D computational models of large organs such as the heart and lungs (Li et al., 2008). These techniques are most effective when the organs are composed of muscle or thick tissue, which allows them to be accurately measured while the subject is alive. However, the interesting features of the oviduct, which are the small internal folds, are made largely of soft tissue. The contrast and detail of models of soft tissue structures captured using medical imaging techniques has been reported to be relatively poor, and insufficient to capture fine detail (Holdsworth and Thornton, 2002; Poddar et al., 2005).

A viable alternative to 3D imaging techniques is through analysis and 3D reconstruction from 2D histology slides (Archip et al., 2006; Qiang et al., 2007). Each image is processed using edge detection to extract a set of cross-sections, and then algorithmically reconstructed. This method is well suited to small scale soft tissue organs, however for organs with complex geometries, several thousand histology images are required to create an accurate model, making this a manually intensive process.

Generating 3D models from cross-sectional data can be split into 4 distinct steps, Correspondence, Tiling, Branching and Surface Fitting. The Correspondence problem relates to determining the topological hierarchy for individual cross-sections. Tiling is the problem of how to connect points of adjacent cross-sections to form an accurate surface. Branching is how to handle the situation when cross-sections split apart or merge together. Surface Fitting is how best to generate a surface for all cross-sections to form a complete mesh. Solutions to these problems have been presented, many of which result in high quality 3D models from cross-sectional or scatter point data (Archip et al., 2006; Meyers et al., 1992; Qiang et al., 2007).

A technique for creating 3D models using sparse cross-sections is described in (Treece et al., 2000). This technique is effective where the region of interest does not change significantly, however for more complex models, a complete, sequential set of cross-sections is required. A method for connecting a simple set of cross-sections together, and interpolating missing sections using b-spline curves is described in (Shreedhara and Indira, 2006), however it is only demonstrated for simple geometries.

These techniques can result in accurate 3D models, however many require highly complex algorithms to be implemented and do not provide structural information, such as the path that a tube follows through the tissue.

Algorithmic approaches have been used to complement biological data, and construct



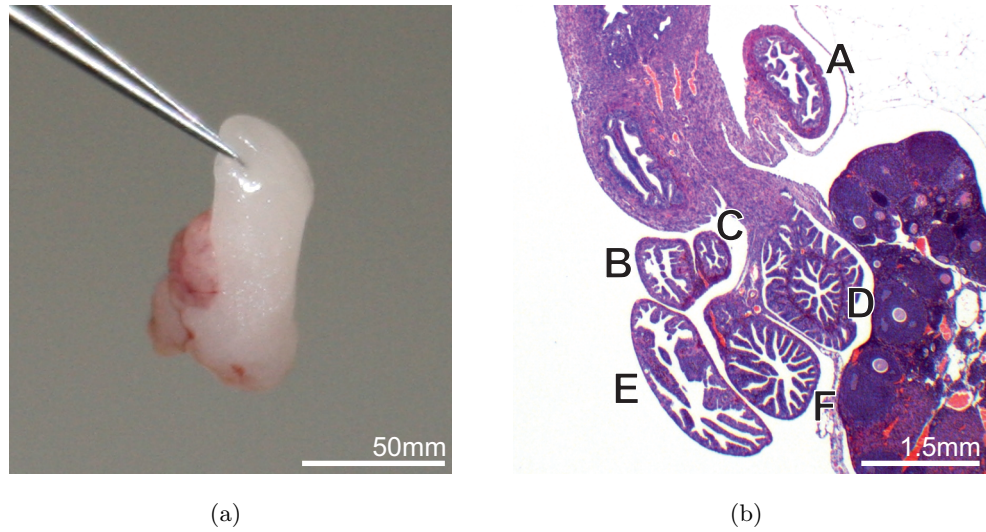
complex pathway based geometries. A fractal like algorithm is described by Kitaoka et al. (1999) to model the airway structure of the lungs in 3D. This model was based on a set of rules determined from the physical properties of the underlying system. The structure consists of interconnected tubes that branch out iteratively, with the thickness, orientation and length based on those of the parent tube. Similar methods have been used to model the flow of blood in the circulatory system (Gabrys et al., 2005), the inter-connections in neuron morphology (Alves et al., 1996), the arterial tree in the placenta (Bergman and Ullberg, 1998) and microvascular networks to aid in the identification of tumours (Craciunescu et al., 1999). 2D imaging data has been used to reconstruct the 3D path followed by coronary arteries (Slager et al., 2000). Smirnov (2003) used a technique inspired by cell division to ‘grow’ the tubular structure of blood vessels.

## 5.2 The structure of the mouse oviduct

As established in Chapter 2, the oviductal environment forms a significant structural obstacle to sperm movement. It is a pre-requisite for all the complex guidance mechanisms currently being investigated and a key factor which could influence the distribution of sperm. Although the oviducts of different species have similar structural features, there are also variations. Within the mouse, the oviductal tube is tightly coiled and naturally encased in fatty tissue, preserving its structure (Figure 5.1(a)). The structure of the mouse oviduct was investigated using histology. A series of histology slides were created from mouse oviducts with a spacing of  $30\mu m$  and a thickness of  $5\mu m$ , as described in Appendix B. In this technique, the tissue is fixed in wax and is then thinly sliced, mounted on a slide and stained. The image can then be captured using a camera mounted on a microscope.

Figure 5.1(b) shows a single histology image taken from the mouse oviduct. This image shows that cross-sections and longitudinal sections from different regions of the same oviduct are visible in the same slide, labelled A - F. This demonstrates the complexity of the tube, and how it folds around itself multiple times. When viewing the individual 2D images, it is not easy to determine which part of the oviductal tube each cross-section belongs to.

The structure of the mouse oviduct changes around the time of ovulation. The UTJ is closed, blocking the isthmic end of the oviduct. When the secretory cells in the walls of the oviduct release their contents, the lumen fills with fluid. The fluid, combined with the presence of the oocyte along with the OCC, causes the ampulla to become distended, with the walls stretching and the folds being pushed into the wall (Hogan



*Figure 5.1:* (a) A photograph of a complete mouse oviduct encased in fatty tissue and (b) a view of a histology image of the mouse oviduct showing its internal structure. Several cross-sections labelled A - F are clearly visible.

et al., 1994; Nilsson and Reinius, 1969).

### 5.2.1 Requirements for model construction

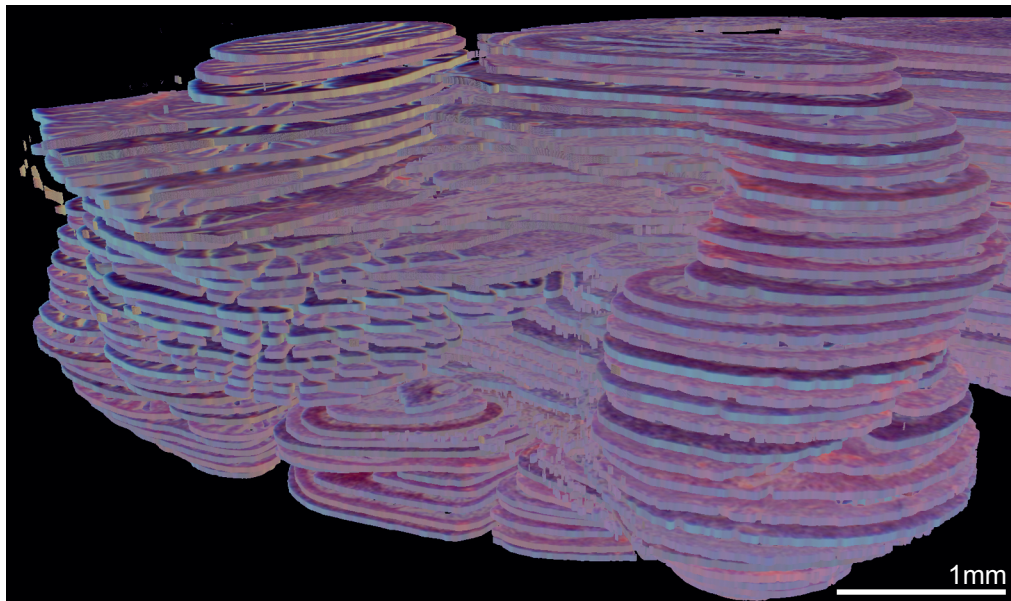
For the 3D models to be useful, they need to be accurately scaled, capture the external bends of the tube and also have some representation of the complexity of the internal folds. Practically, modelling the full details of the internal structure was not possible with the available resources. Including the cellular level detail, such as cilia, secretory cells, crypts and other surface details would result in a massively complex 3D model, which would significantly increase the complexity of the agent model. As the agent model will investigate sperm behaviour around the time of ovulation, the structure of the model should match the conditions at this time period. In order to investigate the influence of the oviduct, it should be possible to isolate the significant structural features of the model in order to determine the influence that each feature has on the agent behaviour. The two main structural features of the mouse oviduct are the internal folds and the external bends.

### 5.2.2 Visualisation of histology images

In order to understand the complexity of the tubal structure, the histology images were visualised in 3D. The images were mapped onto 2D planes and then stacked in

3D space, with spacing relative to their real world size and position. Transparency was then applied to the images to emphasise their 3D structure. As each slice is effectively a 2D plane, it is difficult to view the 3D structure from different angles. Creating a true volumetric reconstruction of the slides is a complex, processor intensive and time consuming process, requiring a more complete set of images for good results.

A simple technique was used to give the illusion of thickness from the static set of images, which makes the 3D structure easy to view from any angle without any pre-processing of the images. Instead of each slice being represented by a single 2D plane, 200 evenly spaced 2D planes with the same image on were stacked close together for each slice. The relative real world distance between the top and bottom plane for each slice was equal to the thickness of the real tissue sample. This technique provides a fast and accurate visualisation of the 3D structure (as shown in Figure 5.2) and makes it easier to determine which part of the oviduct a given cross-section relates to.



*Figure 5.2:* Stacked histology images with transparency applied to the non-tissue regions of each image. The illusion of volume is obtained by closely stacking 200 planes per image slice. The slice thickness is exaggerated from  $5\mu\text{m}$  to  $20\mu\text{m}$  for visual clarity.

### 5.3 Model construction

Common approaches for constructing detailed models of biological structures are complex and computationally intensive, and can result in a highly accurate copy of the real data. However, previous attempts to construct the oviduct of a pig using both Micro-CT scanning (summarised in Appendix A) did not provide a good solution. This was

primarily due to the low resolution of the micro-CT scanner and the inability of the technology to capture the fine soft tissue folds of the internal structure without the presence of contrasting agents. Instead, an algorithmic approach was adopted. This approach was more flexible, and could easily be adapted for different species without the need for large sets of species specific data. The model construction process was split into several steps. Structural information about the histology slides was extracted and linked back to the source images for analysis. The extracted data was then used to parameterise a particle system, resulting in the generation of a set of unique 3D models.

### 5.3.1 Particle systems

Particle systems, which were originally proposed by Reeves (1983) to model objects without solid, well defined boundaries, have since been used to simulate a wide range of phenomena in 3D applications such as fire, smoke, water and other similar special effects. Particle systems have also been used to create 3D objects with solid boundaries. Folds of cloth have been simulated using a particle based 2D mass spring system (Breen, 2000). Szeliski and Tonnesen (1992) describe a technique for constructing dynamic surfaces by self-orienting particles, which form surfaces instead of solid objects.

Particle systems have also been used to model biological systems. Nedel and Thalmann (1998) used a particle system in combination with a 2D mass spring system to model the elastic properties of muscle tissue. Stahl et al. (2002) used a particle system to model the surface of deformable objects with elastic properties, which has been used to simulate the deformation of biological organs such as the heart. Amrani and Shariat (2000) use multiple layers of particles with different properties to provide a more realistic simulation of 3D tissue.

### 5.3.2 High performance computing

Due to the complexity of the oviductal environment, the massively parallel processing power of the Graphics Processing Unit (GPU) and multi-core Central Processing Units (CPUs) have been used to perform the image analysis and to run the particle simulation. There are a number of techniques available for programming on the GPU, however Compute Unified Device Architecture (CUDA), which is a set of extensions for the C programming language, has been widely adopted by the scientific community (Kirk and Hwu, 2010; Sanders and Kandrot, 2010). Unlike the CPU, where data and instructions are processed serially, the GPU runs in parallel, applying the same set of instructions to multiple elements of a data set at the same time. These differences in architecture

require programs on the GPU to be written in a very different way to programs on the CPU.

At the lowest level, the GPU is able to process a large collection of threads concurrently, with each thread mapping onto a single element in the input data set. The input data set is sub-divided into blocks, which are a 1D, 2D or 3D grouping of threads, with each block processing a specific number of elements. All threads within a single block are processed cooperatively and can interact with each other using shared memory. Synchronisation points can be defined to ensure that all threads within the block are at the same point of execution. Blocks are grouped together into a grid, which is a 1D or 2D grouping of blocks. The GPU also has several different types of memory, each having a different scope, speed and usage. The choice of block size and the type memory used can have a significant impact on the performance of the program (Kirk and Hwu, 2010).

Force based particle systems have been traditionally implemented on the CPU, however, more recent implementations have been modified to run on the GPU (Kruger et al., 2005). Several recent studies have demonstrated high performance, real time processing using the GPU for physics simulations and other computationally intensive calculations (Crane et al., 2007; Harada, 2007). A force based N-Body simulation has been described that represents 16,384 particles and runs over 50 times faster than a comparable CPU implementation (Nyland et al., 2007). A real time rigid body collision simulation with 16,384 interacting objects was implemented on the GPU, running in real time (Harada, 2007).

## 5.4 Extraction of structural information from histology images

Although the complex 3D structure can be clearly seen by visualising the stacked histology images, this does not provide structural information for use in predictive simulations. This section describes the process used to identify the individual tubular cross-sections from each image, and connect those cross-sections together to determine the path that the oviduct follows in 3D space.

### 5.4.1 Identification of lumen cross-sections

The first task was to identify the individual tubular cross-sections in each image. The images were processed using a combination of CUDA on the GPU and multi-core CPU processing. For each pixel in an image, the pixel intensity is taken to be the average colour intensity of the 3 by 3 region surrounding the pixel. If this value is smaller than a given threshold, then the pixel is determined to be part of the tissue. Using

the neighbourhood average instead of the individual pixel value when determining the tissue regions helps to reduce the impact of anomalies in the image caused by the histology process. During this process, all pixels which represent a boundary between the tissue and non-tissue regions of the image are also identified.

The next step was to identify the cross-sectional lumen, which is the space inside the tube. This operation combined a flood fill of the white space surrounding the tissue, and identification and grouping of all pixels within each area of lumen in a single step. Each pixel in the image was assigned a unique sequential id, matching its index in a 1D array. The pixels at the edge of the image were assigned an index of 0. The image was then mapped onto blocks on the GPU using a size of 32 by 16. Within the kernel, an area of shared memory with a 1 pixel buffer of size 34 by 18, was then created to store the current id of each pixel within the block, and the pixels surrounding this portion of the image.

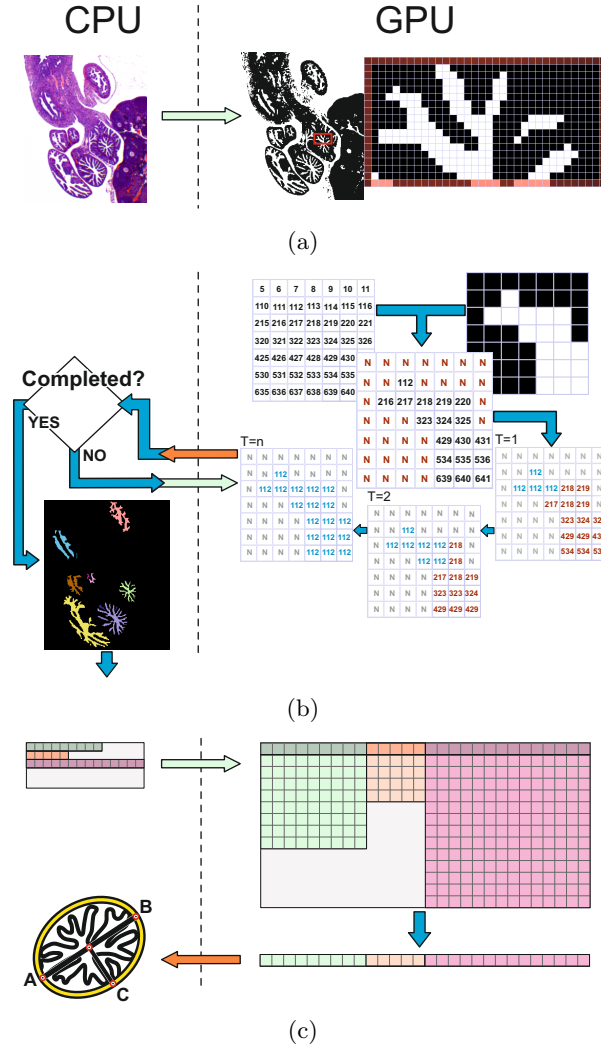
The shared memory was populated by the corresponding pixel using a combination of the id and the binary tissue image. Where a pixel was determined to contain tissue, the id was assigned a value of N, where N is the total number of pixels in the image, otherwise the id remained unchanged. Next, the pixels adjacent to the buffer region populated the adjacent values into the shared memory array. The threads within the grid were then synchronised. This process is shown in Figure 5.3(a).

The second part of the kernel evaluates the id value of each non-tissue pixel in relation to its immediate neighbours using a 1 pixel radius and assigns the current pixel to the lowest id of all neighbours. This was initially implemented as a nested loop, however it was more efficient to explicitly write out the evaluation as 8 nested  $\min(X, Y)$  functions.

The threads were then synchronised, and the evaluation repeated until none of the pixels from any of the blocks change. If the value of any pixel from any block changes id in any of the evaluation iterations, then a global flag is set.

This global flag is read back to the CPU, and if the flag is set, the data is processed again. This is repeated until none of the ids change. Once complete, every pixel within every area of the image which is enclosed by tissue holds the same id number, with different areas having different ids. As the outer border of the image was assigned an id of 0, all white space surrounding the tissue also has an id of 0. This process is shown in Figure 5.3(b).

From this, each pixel within the identified lumen regions is assigned a neighbourhood intensity value. This is determined as the number of pixels within a 10 pixel radius which hold the same id value. This relates to the clustering of space within the lumen.



*Figure 5.3:* The complete cross-section identification process for a single image and how the different processes are distributed between the CPU and GPU. (a) The image is transferred to the GPU, where the tissue regions and boundaries are identified. The red border represents the 1 pixel region adjacent to the block. (b) The non-tissue regions and the id associated with each pixel are combined, then the remaining ids are iteratively reduced to hold the minimum value of all neighbouring pixels. Once completed, the data is transferred back to the CPU. The image shows a representation of the data with each unique remaining id holding a different colour. (c) The identified regions are combined with the border data and converted into a set of lists, one for each cross-section and passed to the GPU. Within each list, N-Body comparison and 1D reduction is performed to determine the largest distance between two points within a cross-section. This is then transferred back to the CPU and used to calculate an ellipse which encapsulates the corresponding cross-section.

Regions with large gaps receive a high intensity and regions with small gaps have a lower intensity. This helps to identify the correct midpoint of all cross sections. It also gives an indication of the size of the space within each cross section, and can be used to identify abnormal luminal spaces, which is where the histology slide passes through the top or bottom of a tube, and two cross sections are starting to join together.

The common pixel index information was then streamed back to the CPU, and combined with the previously identified boundary information into a set of lists, with each list containing the indexes of all boundary pixels with a common id. Lists smaller than a specific threshold value were then removed, helping to ensure that small holes in the tissue are excluded from further processing. As this process is sequential, it is easily and optimally performed on the CPU. Transferring data between CPU and GPU is often considered undesirable due to the limited bandwidth. However, when the processing time of a set of operations becomes greater than a second, the transfer time becomes negligible in comparison. Although it could have been implemented on the GPU, it was far quicker and simpler to transfer the data to the CPU to perform the serial process, then transfer the restructured data back again.

The linear lists of indices were then passed to the GPU as a 1D list, along with a second list containing the cumulative size of each list. The data was then processed one list at a time on the GPU to determine the maximum distance between all points within the same region. This is effectively an N-Body comparison, with the position of every point being compared with the position of every other point within a single list. For each list, the points are mapped onto a 2D grid (X, Y) where the lengths of (X) and (Y) are equal to the number of points in a list (N).

Two arrays of shared memory, one to store the distance and one to store the corresponding index were created. For each point in X, the distance between the corresponding point in Y was stored in the shared memory. The threads in each block were then synchronised. For every point in (X), all associated records in shared memory were iteratively compared using standard reduction techniques (Kirk and Hwu, 2010), selecting the pair with the greatest distance between them. This was then written to global memory, with each point having a corresponding distance record for every block in the (Y) dimension. A second reduction was applied to the global memory, resulting in a single distance record point, as shown in Figure 5.3(c).

This information was then transferred back to the CPU and used to determine an ellipse which encapsulates the lumen. Figure 5.3(c) shows how the elliptical region was calculated. The two points (A and B) furthest apart form the major axis of an ellipse. The true midpoint of the cross section was then determined using the



previously calculated neighbourhood intensity. The midpoint was determined to be the weighted average of the X and Y positions of all pixels within the cross section, where the neighbourhood intensity of each pixel was the weighting. This means that the midpoint was closer to pixels with a higher intensity. The identified major elliptical axis was then offset so that the centre of the line coincided with the newly calculated midpoint.

The two points perpendicular to the major axis crossing the mid-point with a length equal to the length of the major axis were determined, then the nearest two points within the region to each perpendicular point were identified. The minor axis (C) was determined by a line with double the length of the point furthest away from the major axis which ran perpendicular to the major axis crossing the midpoint.

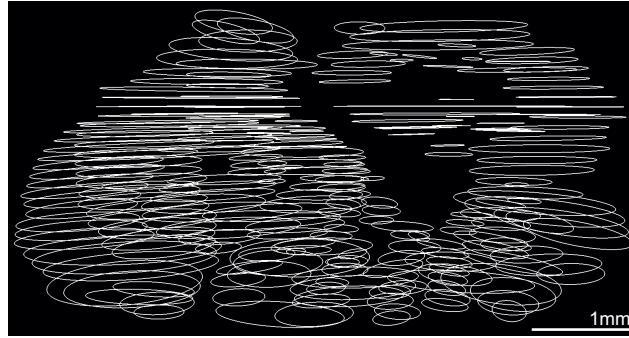
After the elliptical region was defined, it was then analysed to determine if it was a normal region or an abnormal region. An abnormal region was identified when the ratio between the elliptical length and width was larger than a specified threshold, whereby the ellipse is long and thin. Once identified, an abnormal region was split into three sub regions, each with its own elliptical region. The major elliptical axis was split into three points, and from each point a new elliptical region was calculated using half the parent long axis length and its own long axis length. The midpoint for each sub-region was then recalculated using the neighbourhood intensity of all pixels covered by the area covered by the sub-region, and the remaining elliptical region properties were determined as before.

All images were processed, resulting in a set of stacked elliptical regions in 3D space which encapsulate the lumen of different cross-sections, as shown in Figure 5.4(a).

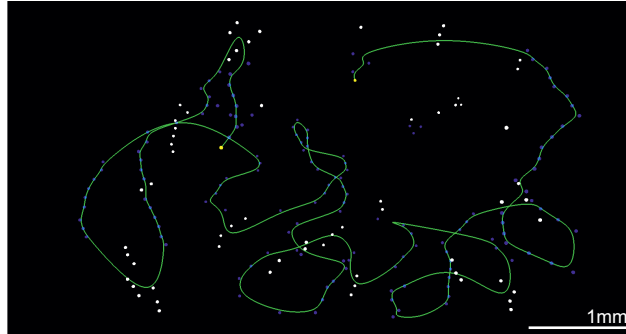
#### 5.4.2 Curved path identification

Existing solutions to the correspondence problem, which relates to the topology of cross-sections, are more complex than needed for this application. Also, due to the nature of the oviductal tube, with multiple sections in a single region of tissue, existing techniques would not result in correct topological ordering of the cross-sections.

Instead, a simple CPU based algorithm was developed to connect all elliptical regions together to recreate the 3D path that the oviduct tube follows. Initially, a fully automated algorithm was developed incorporating domain specific knowledge about the structure of the tube. However, this approach was found to be too limited when applied to multiple data sets, and was unable to reliably identify the correct path through the oviduct images without a more complete set of data. Instead, a semi-automated



(a)



(b)

*Figure 5.4:* The relationship between (a) the elliptical regions identified from the histology images and (b) the final 3D curve after the path has been identified.

method was developed, which still incorporates knowledge about the structure of the tube to simplify the algorithm.

A modified tree structure was implemented to store the connection structure. This node based structure is similar to a traditional tree, and an elliptical region is associated as a property of each node. However, unlike a traditional tree structure, each node can have a web of multiple connections. Each connection also has additional information associated describing the nature of the connection. The additional information can contain both relative positional information, such as if the connected node is above, below or lateral to the current node, and confidence related information, such as whether the connection is considered to be a primary connection or a secondary connection.

Histology images are not always perfect, some artefacts in the slides, such as holes in the tissue, are identified as cross-sections, and some true cross-sections are ignored, due to tears in the tissue during slicing. This can result in gaps in the structure. Before all potential connections are established, the algorithm attempts to generate new potential

elliptical regions to fill these holes based on knowledge about how the tube is formed. If an elliptical region exists in one slide, and an overlapping region exists two or three slides down, then new elliptical regions are inserted in the gaps, with the new region being the interpolated average of the two existing regions, as shown in Figure 5.5(a).

Next, all potential connections were established between all elliptical regions. Two regions were deemed to be primary connections if they were above or below the one another and their midpoints were contained within each others elliptical regions. Two regions were deemed to be secondary connections if they were above or below one another and their elliptical regions partially overlapped. Any intersection with an interpolated elliptical region was marked as a secondary connection.

For elliptical regions which were identified as abnormal regions, intersections were performed with the sub regions using the same rules. This means that a primary connection was established if the elliptical region intersects with the midpoints of any of the sub regions. Some of the source image data sets were missing some images from the very top and bottom of the tissue. In instances where two disconnected tubes in close proximity end at the top or bottom of the image set, secondary lateral connections were created between these two regions.

A set of simple functions was then applied to reduce the number of connections. Clustered groups of nodes, which are branched off from the main path were merged together. These typically occurred at the top of bottom of a bend, where multiple cross-sections were identified for a single section of lumen. Single strands which branched off the main path and did not re-connect were also merged. These also typically occurred at the top of bottom of bends, or where artefacts in the histology images caused the false identification of a cross section. Where two strings of primary connected nodes were interconnected by secondary connections, these surplus secondary connections were removed.

These functions were used to apply bulk modifications to the path. After applying these functions, the complete path was visually apparent, however multiple surplus connections still existed. Due to the close and complex nature of the bends, and the need for this approach to be accurate, surplus connections were manually identified and removed using a simple visual interface. Four simple operations were performed as necessary, which were (a) removing a connection between two nodes, (b) adding a connection between two nodes, (c) removing all connections to a specific node and (d) removing all nodes connected together in an isolated cluster. Once the path was complete, the end of the path representing the isthmus region was manually identified.

For a fully automated solution to be developed, a large number of data sets would be required to train the system and determine a comprehensive set of connection rules. As the main aim of this research was the creation of a set of 3D models, and not the development of a fully automated solution, the small amount of time taken to remove the surplus connections is justified. This approach worked well for multiple data sets, even when there are large gaps between images.

Once a single path had been identified, the points along the path were then connected together using b-spline curves to form a smooth path, which allowed an estimation of the physical length of the oviduct tube to be determined. Using the curve points as the control points in the b-spline algorithm resulted in a reduction in the length of the tube. Instead, the control points  $A$  and  $B$  for each point along the curve were derived from

$$A = Pc + (Vp \times Dp \times k) \quad (5.1)$$

$$B = Pc + (Vn \times Dn \times k) \quad (5.2)$$

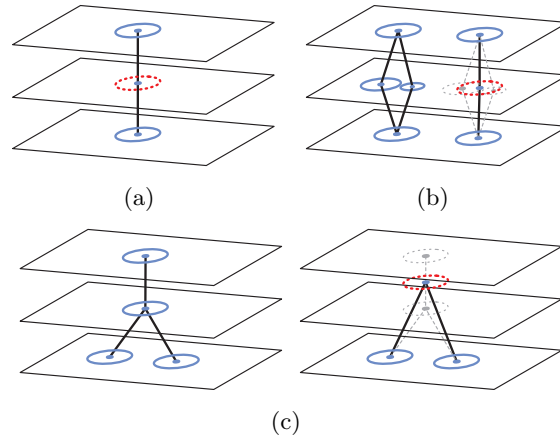
where  $Pc$  is the current point along the curve,  $Vp$  and  $Vn$  are the vector between the current and previous point in the curve and current and next point in the curve respectively,  $Dp$  and  $Dn$  are the distance between the current and previous point in the curve and the current and next point in the curve respectively, and  $k$  is a constant. This resulted in a closer matching of the smoothed curve to the control points. Figure 5.4(b) shows the final curve extracted from the cross-sections. Finally, a set of evenly spaced points along the curve were extracted.

### 5.4.3 Internal cross-section generation

The points at which the curve crossed each image was identified, and a series of sequential images were generated, combining the original images with the curve crossing information. This set of images related each 2D cross-section with its corresponding 3D position along the length of the oviduct. A set of measurements were taken for each image, which are summarised in Table 5.1. How each measurement was taken is shown diagrammatically in Figure 5.6.

These measurements were taken from three different images sets, with all three showing remarkable similarities along the length. The averaged measurements for all three images sets were taken at regular intervals along the length of the oviduct tube.

As described in Section 5.2 the structure of the mouse oviduct changes around the



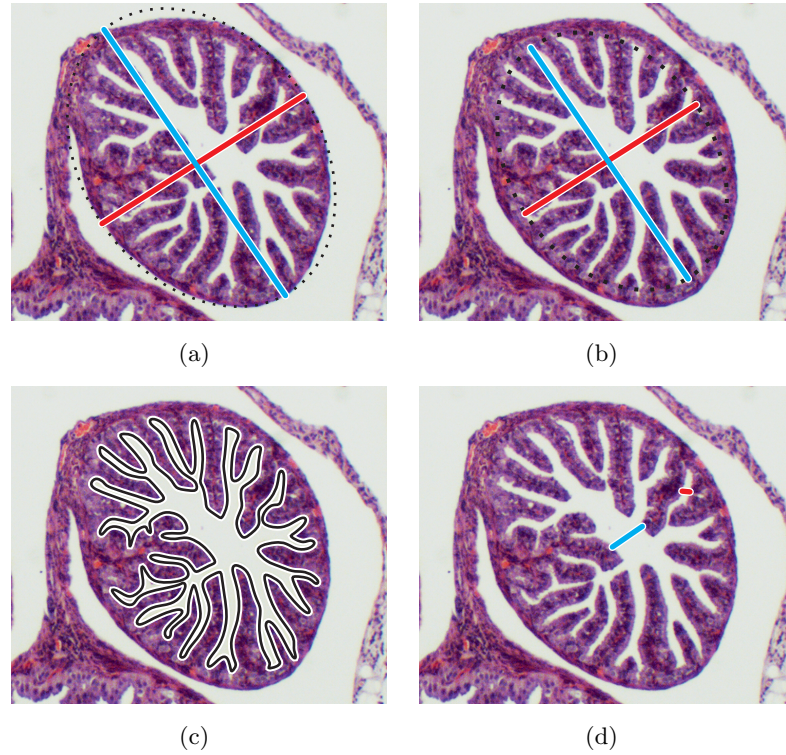
*Figure 5.5:* The path connection rules used to determine the cross-sectional topology. (a) A missing cross-section is identified in between two vertically overlapping cross-sections. (b) Clusters of points which are interconnected are merged together. (c) A single cross-section above or below two merging branches (left) is assumed to be the top or bottom of a bend in the tube, so they are merged together (right).

Measurement	Description
Outside Wide	The widest diameter of the outer tube
Outside Narrow	The narrowest diameter of the outer tube
Inside Wide	The widest diameter encapsulating the lumen
Inside Narrow	The narrowest diameter encapsulating the lumen
Wall Length	The length of a line drawn along the internal wall of the oviduct
Wide Lumen	The width of the lumen at the widest point
Narrow Lumen	The width of the lumen at the narrowest point

*Table 5.1:* Describes the measurements taken from the histology image cross-sections.

time of ovulation. As the ampulla region of the oviduct fills with fluid and eggs it becomes swollen, and the internal folds become pressed into the wall. The histology images in the source dataset did not have this alteration. In order to compensate for these changes, the measurements taken from the histology slides were modified. The outer and inner length and width and lumen measurements were increased to match the description in literature (Nilsson and Reinius, 1969). The final measurements are shown in Appendix C.

The modified measurements were then used to computationally generate a set of linear 2D cross-sections of the internal and external structure of the oviduct, based on the techniques described in (Burkitt et al., 2010). The techniques use a force based



*Figure 5.6:* The different measurements taken from the histology images. (a) The outer wide (blue) and narrow (red) diameter of the tube. (b) The inner wide (blue) and narrow (red) diameter of the lumen. (c) The length of the internal lumen. (d) Measurements from the wide (blue) and narrow (red) portions of the lumen.

particle system which models the internal folds and external wall of the oviduct as two single loops of connected 2D particles. The particles move over time with respect to the forces acting on them, causing the structure to deform. The number of particles also increases over time, allowing the increase in complexity of the internal structure from one end to the other to be represented. The position of all particles in each loop, which form a single 2D cross-section, are then recorded at regular intervals. Once complete, the 2D cross-sections are then connected together to form a complete 3D model. This results in a unique, accurately scaled 3D model of the oviduct. However, the model is completely straight, and does not represent the complex bends which exist in the real organ. The system described in (Burkitt et al., 2010) is accelerated on the GPU using C for Graphics (Cg) shaders, which are primarily used for producing visual effects in 3D graphics. However, this implementation was considered complex and limited, so was rewritten to use CUDA. Figure 5.7 shows a selection of computationally generated cross sections for different regions of the oviduct. Validation of these cross sections is discussed in Section 5.6.

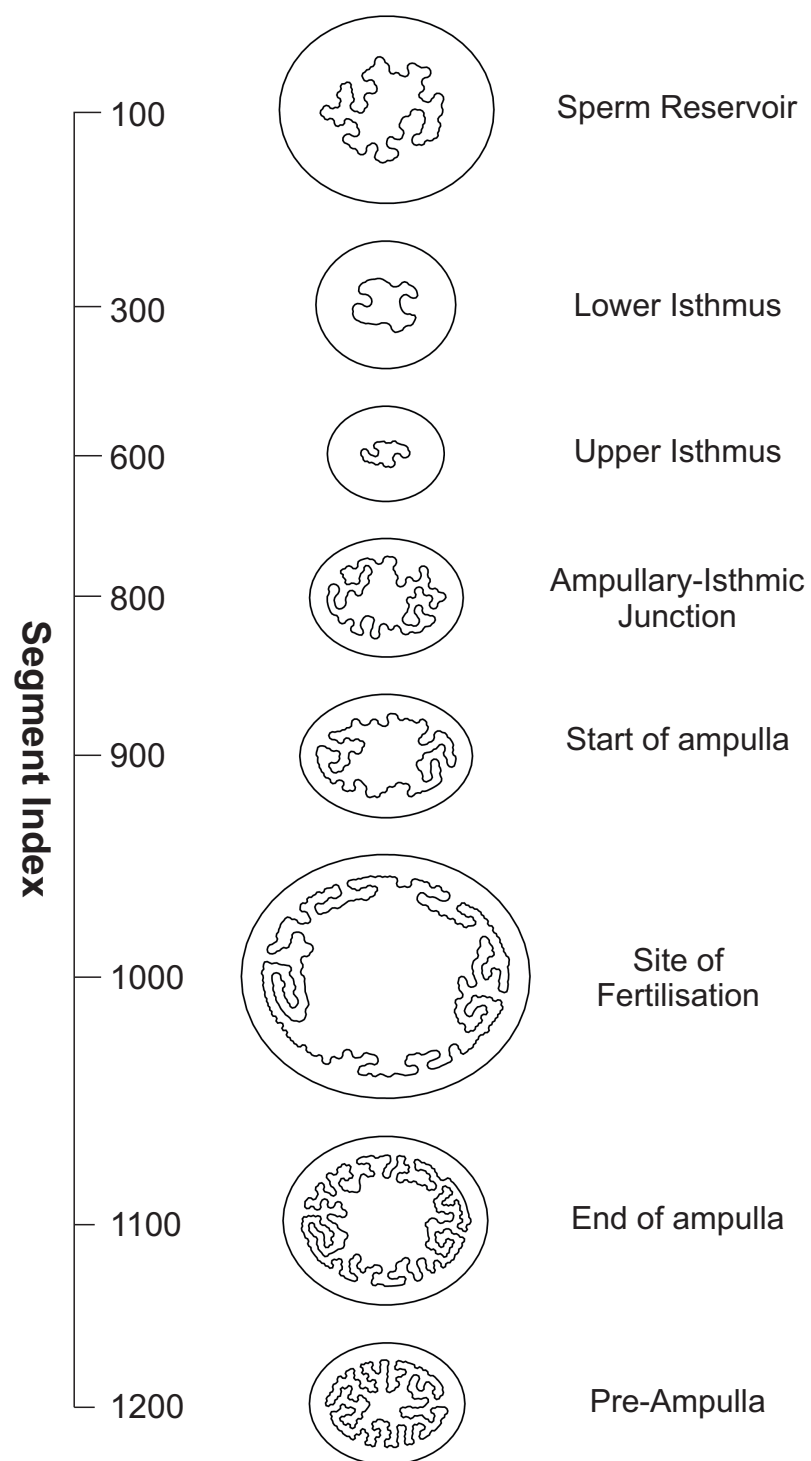
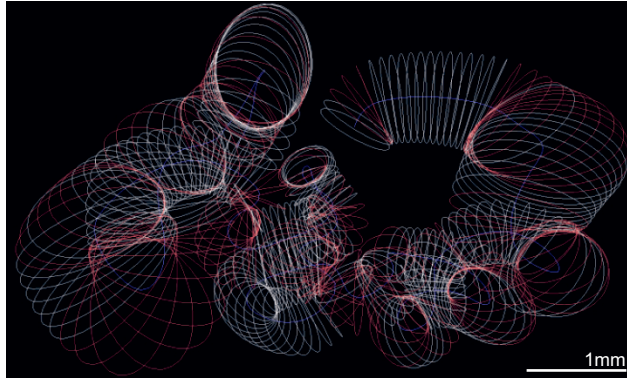


Figure 5.7: A series of generated cross-sections and how they relate to different regions of the oviduct.

### 5.5 Binding the cross-sections to the 3D curve

The 3D curve has a complex form containing tight bends which pass close by each other in parts. Initially, the computer generated 2D cross-sections were bound to matching points along the curve, and oriented to follow the curve gradient. However, this resulted in self intersections and crossover of adjacent sections, as shown in Figure 5.8. Several techniques were used to try to resolve these intersections. Overlapping cross-sections were rotated about their mid-points until they no longer overlap, which resulted in sharp angles between adjacent sections in some parts of the model and unwanted deformations in others. Adjusting these techniques still failed to solve the problem, and did not resolve self intersections.



*Figure 5.8:* Cross-sections bound to the 3D curve and rotated to match the curve gradient. Collisions between adjacent sections are red.

The final successful approach was to resolve the collisions using a particle physics based simulation. A series of particles, which are points in 3D space, were connected together in a closed loop to describe the elliptical boundary of each 2D cross section. The speed and position of the particles are modified over time by a set of virtual forces. The forces between particles maintain the shape of the cross-section and defines how the particles in different cross sections should interact. Each cross-section could have been modelled as a static body, however this would have required several complex calculations. Typically, particle physics based simulations use a set of forces which act on particles individually. This simulation differs slightly with the introduction of the concept of a particle group. This allows a force function to either process an individual particle or on entire groups.

The standard Runge-Kutta 4 (RK4) method was used for the numerical solution of the particle-force resolution for all implementations. This method involves four steps



of calculations. Before each step, the forces acting on each particle are calculated and then used to calculate the new position and velocity of all particles within the system. This process is then repeated 3 times. With each step, the cumulative force acting on each particle and its resulting velocity is calculated. With the final step, the final position and velocity of each particle is then calculated based on position at the final step and the cumulative force and velocity. All forces are implemented as individual functions, as is each step in the RK4 method.

### 5.5.1 Force definitions

A set of forces was defined to help orient the cross-sections in a way that prevented self intersections, maintained the integrity of the model and followed the 3D path.

#### 5.5.1.1 Base forces

A set of forces upon which other forces are based is described here.

A spring force  $F_s$ , which is based on Hooke's law of elasticity, is used to maintain a connection between two adjacent particles. This force is defined as

$$F_s = -((k \times (x - r)) + (c \times v)) \quad (5.3)$$

where  $k$  is the spring strength,  $x$  is the current distance between the two particles,  $r$  is the rest length (the distance between particles where no force is applied),  $c$  is the damping strength, and  $v$  the current difference in velocity between the connected particles.

A dynamic spring force is a variant of the base spring force. In a slight modification to the spring described above, the length of the spring is also dynamic and changes over time based on the following formula

$$r_{new} = (x \times \delta) + (r \times (1 - \delta)) \quad (5.4)$$

where  $r_{new}$  is the newly calculated rest length,  $x$  is the current distance between the two particles,  $r$  is the old rest length and  $\delta$  is an adjustment factor, with a range of  $[0, 1]$ .

An anchored force is another variation of the base spring force, where instead of two particles being connected, a single particle is connected to a fixed point in space.

A simple frictional force  $Fr$  which is applied as a constant damping force to all particles is defined as

$$Fr = -v \times k \quad (5.5)$$

where  $v$  is the current velocity of the particle and  $k$  is the friction strength, which has a range of  $[0, 1]$ . This serves to prevent forces from becoming too large, and destabilising the simulation.

### 5.5.1.2 Cross-sectional integrity

Two different forces help to maintain the integrity of the cross-section. By integrity, we mean maintaining the shape and size of the cross-section, and helping to ensure that all points lie on or close to the same plane.

The Spoke Force (Figure 5.9(a)) is a high strength, high damping spring force which maintains the basic structure by ensuring all particles within a cross-section maintain the same distance from the centre.

The Radial Integrity Force (Figure 5.9(b)) Is also a high strength, high damping spring force which maintains the distance between adjacent particles, and also maintains the distance between opposing particles. The magnitude of these forces is larger than other forces used within the system, which helps to ensure that the integrity of the cross-section is maintained.

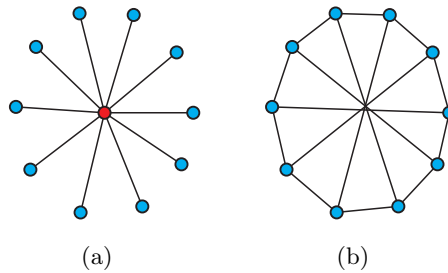


Figure 5.9: (a) The Spoke and (b) the Radial Integrity Forces which maintain the radial shape of the cross-section and the connection between adjacent and opposite particles.

### 5.5.1.3 Structural integrity

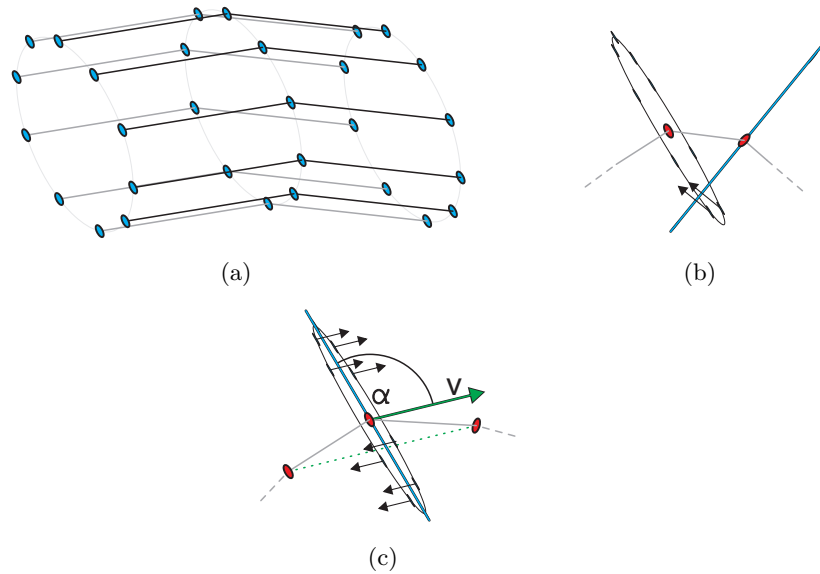
The structural integrity of the tube is enforced using a combination of three different forces.

The Loose Integrity Force (Figure 5.10(a)) is a low strength, high damping dynamic spring with a low adjustment factor which connects particles from adjacent cross-

sections together. This helps to form smooth curves when the cross-sections go around a bend, as the rest length for connections on the outside of the bend are made longer, the rest length for connections on the inside of the bend are made shorter.

The Crossover Force (Figure 5.10(b)) helps to resolve adjacent cross-sections which intersect each other. The particles from one cross-section are processed in relation to the plane formed by the adjacent cross-section. If the particles are in front of the plane, then a force is applied to them to move in the direction opposite to the plane normal.

The Gradient Force (Figure 5.10(c)) prevents the tube from collapsing. The angle between the mid-point of each cross-section and each particle around the cross-section ( $\alpha$ ) and the vector between the mid-points of the previous and next cross-sections ( $V$ ) is calculated. If the angle is too far away from 90 degrees, A force is applied to all particles either in the direction of ( $V$ ) or opposite to ( $V$ ), to reorient the cross-section back towards a 90 degree angle. Without this force, the tube tended to become flattened in parts.



*Figure 5.10:* The different Structural Integrity Forces. The Loose Integrity Force (a) maintains a connection between particles from adjacent sections. The Crossover Force (b) prevents intersections between adjacent sections. The Gradient Force (c) prevents the tube from flattening.

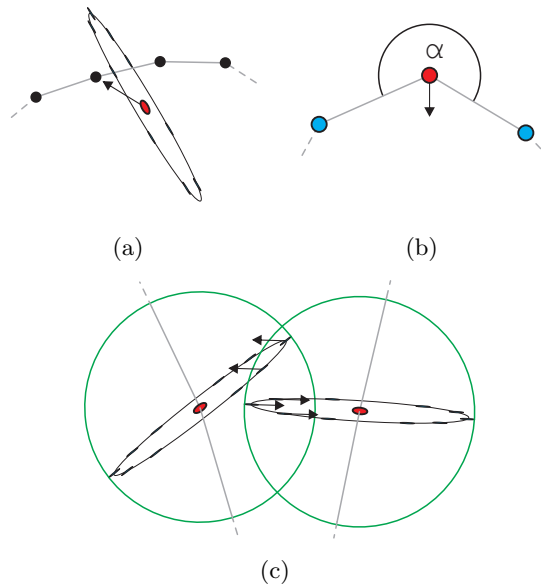
#### 5.5.1.4 Form and realism

The final set of forces help to constrain the cross-sections to the 3D curve while maintaining the physical realism of the structure.

The Position Force (Figure 5.11(a)) maintains a low strength, low damping anchored spring link to the original position along the curve of the cross-section. This helps to ensure that the simulation binds the cross-sections as closely as possible to the curve.

The Smoothing Force (Figure 5.11(b)) helps to smooth out and sharp edges, resulting in a more natural structure. For each particle, if the angle between the previous and next particles is less than a specific threshold, a forces is applied to the particle to move it towards the two adjacent particles, reducing the angle.

The Collision Force (Figure 5.11(c)) prevents the tube from self intersecting. For every cross-section, the radius between the mid-point and the furthest particle away from the mid-point is calculated. All cross-sections are then compared, allowing a buffer of 8 sections either side to prevent adjacent sections from being evaluated. If two radii intersect, then all particles from each cross-section are tested to see if they are within the radius of the other cross-section. If they are, a force is applied to each to push it away from the intersected cross-section.



*Figure 5.11:* The Form and Realism Forces. The Position Force (a) ensures that the mid-point remains close to its corresponding position along the curve. The Smoothing Force (b) removes sharp angles by applying a force to the middle particle when the angle becomes large. The Collision Force (c) prevents different sections of the tube from intersecting.

The Friction Force acts to reduce the magnitude of the forces building up in the system. Due to the strength of some of the forces, a high amount of friction is used in the system, reducing the velocity of all particles by 100% every iteration. This also helps

to maintain a tight structure and ensures that the particles only move as far as they need to.

### 5.5.2 Force resolution computation

As the simulation progresses, the particles shift their position until all collisions and self intersections are resolved, and the structure is gradually smoothed. The simulation runs until it reaches equilibrium, which is established once the velocity of all particles remains less than  $0.25\mu m$  per iteration for 100 consecutive iterations.

Once the simulation reaches equilibrium, the translation and rotation of each cross-section is calculated and stored as a set of matrix transformations. These matrix transformations are then applied to the original linear 2D cross-sections to create the complete 3D model, as shown in Figure 5.12.

### 5.5.3 Implementation

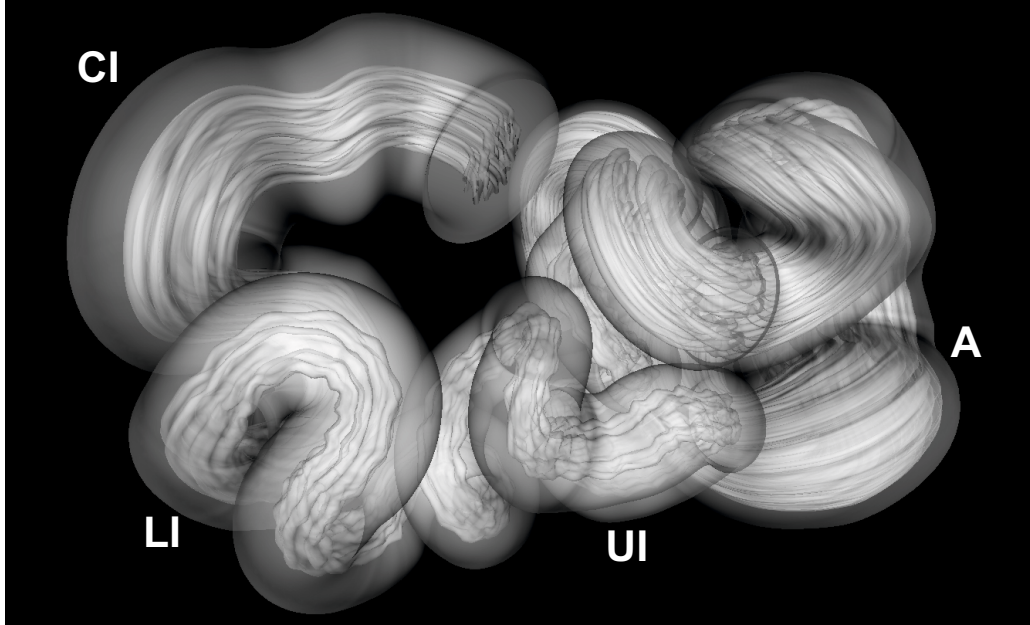
The particle physics based simulation was initially implemented on the CPU using a single processor. This was soon modified to make use of multi-core programming techniques by sharing the simulation processing between all the CPU cores on the computer, resulting in a faster simulation. However, with a complete set of input data (301 cross-sections with 12,341 particles), the solution took 86 minutes to reach equilibrium. The simulation was then ported to run on the GPU using CUDA. This required several changes to the initial implementation to make best use of the GPU processor, but reduced the total simulation time to around 7 minutes.

#### 5.5.3.1 Data structures

On the CPU, the vector data was stored in continuous arrays of single precision floating point Vector3 structures, with an X, Y and Z component. On the GPU, the data could be structured in two different ways. CUDA does have inherent support for vector types, however due to the way memory is accessed, coalescing is only possible using 2 or 4 component vectors. As our data is stored in a 3 component structure, this means we would have to use 4 component vectors, with redundant data in the 4th component. The alternative to this is to store the data serially, in a 1D array, with all the X values first, then the Y values then the Z values.

Using the 4 component vector would result in a 25% increase in the size of the data stored on the GPU. There would also be an unnecessary increase in memory bandwidth usage with 4 component vectors over single values, which could have a negative impact

on performance. Both versions of the data structure were implemented for comparison.



*Figure 5.12:* A translucent view of the completed 3D model after all collisions have been resolved. (CI) is the caudal isthmus just above the UTJ. (LI) is the lower isthmus. (UI) is the upper isthmus. (A) is the ampulla region.

### 5.5.3.2 Force functions

For all implementations, each force and each step in the RK4 method were implemented as individual functions. The functions which act on particles within the system show the most distinctive differences between implementations. When run on a single-core CPU, each function is applied to every particle [0 n] one at a time. For the initial naive multi-core implementation, each particle was processed by an individual CPU thread, and each function was applied to each particle within its own thread. This resulted in a large number of CPU threads being created and disposed, and although thread pooling was adopted, the increase in performance was minimal. This naive approach was improved by creating a single thread per CPU core, and evenly distributing the particles between threads. Each function was then applied to all particles within each thread. Multi-threading was implemented using the built in Thread and ThreadPool classes in the .Net programming language.

Due to the capabilities of the GPU to process a large number of threads simultaneously, the optimal solution was to process each particle on its own GPU thread. This means that the naive implementation for a multi-core CPU is the optimal solution for

the GPU. This is primarily due to the fact that on the GPU, thread switching is performed in hardware, whereas on the CPU it is performed by the underlying operating system, and thus has a much higher cost. The way a function is applied to the different implementation types is shown in Figure 5.13.

	Single CPU	Multiple CPU CPU CPU CPU				Single GPU
<b>F0</b>	[0..n]	[0..a]	[a..b]	[b..c]	[c..n]	[0] [1] [2] [3] [..] [n]
		----- Sync -----				----- Sync -----
<b>F1</b>	[0..n]	[0..a]	[a..b]	[b..c]	[c..n]	[0] [1] [2] [3] [..] [n]
		----- Sync -----				----- Sync -----
		...				
<b>Fn</b>	[0..n]	[0..a]	[a..b]	[b..c]	[c..n]	[0] [1] [2] [3] [..] [n]
		----- Sync -----				----- Sync -----

Figure 5.13: Differences between how functions [F0 Fn] are processed on single-core, multi-core and GPU processors.

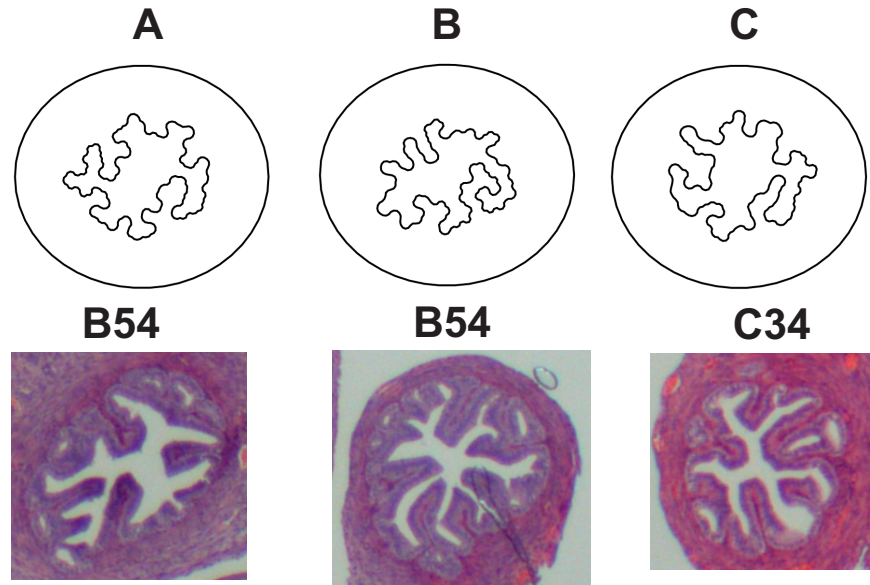
## 5.6 Validation

This section describes how the internal folds and external bends of the 3D model have been validated.

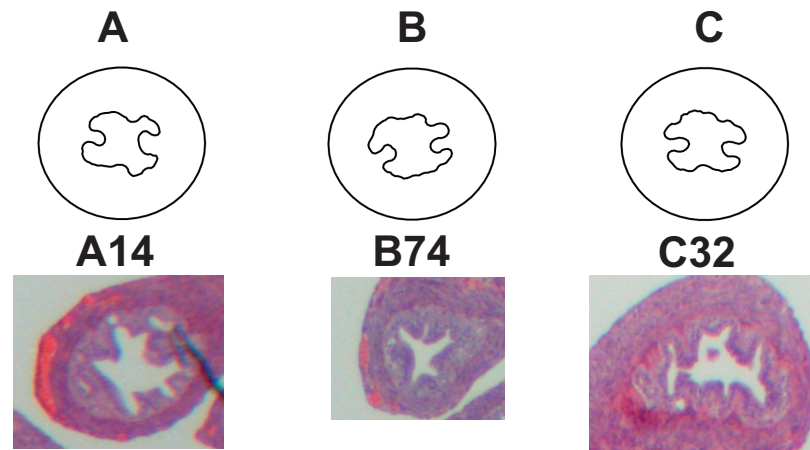
### 5.6.1 Internal cross section validation

Validation of the internal tissue folds was performed by visually comparing the generated cross sections with the histology images at corresponding positions along the oviduct. The comparisons were performed with regard to the scale of cross section and features within each image and the complexity of branching. Figure 5.14 shows a comparison between the generated cross sections of the sperm reservoir and the histology images. The lumen of the histology images has a similar size to the generated images. The complexity of branching is also similar, although the generated cross sections have slightly shorter branches and a larger open region in the middle of the lumen.

Figure 5.15 shows a comparison between the generated cross sections of the lower isthmus and the histology images. The complexity of the cross sectional images in the lower isthmus is far simpler than in the sperm reservoir, and the tube is far narrower. The complexity of branches in the generated cross sections and the histology images is similar, and the size of the lumen is also similar.



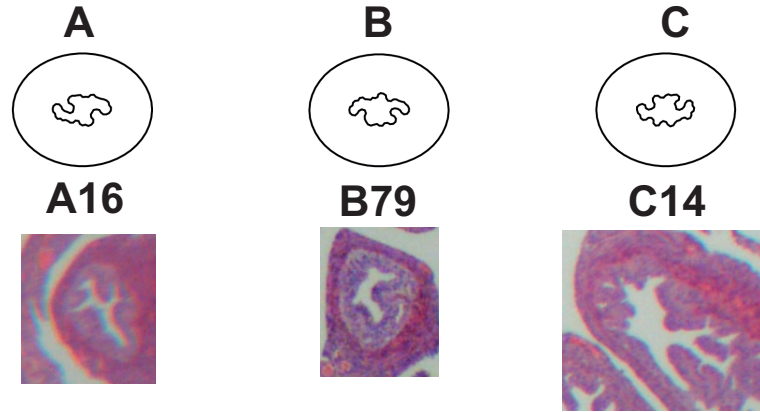
*Figure 5.14:* Comparison between three generated cross sections A, B & C and three histology images from the corresponding position in the sperm reservoir of the oviduct. Two images from histology set B slice 54 and one image from histology set C slice 34 are shown for comparison.



*Figure 5.15:* Comparison between three generated cross sections A, B & C and three histology images from the corresponding position in the lower isthmus of the oviduct. One image from histology set A slice 14, one image from histology set B slice 74 and one image from histology set C slice 32 are shown for comparison.



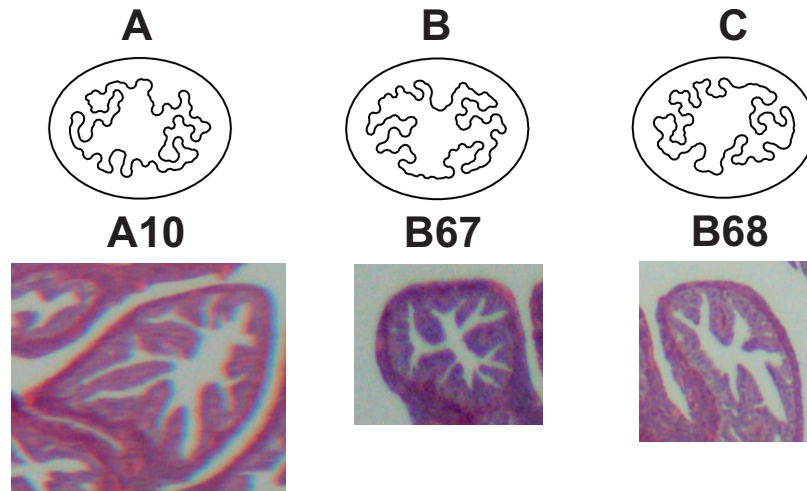
Figure 5.16 shows a comparison between the generated cross sections of the upper isthmus and the histology images. The lumen of the upper isthmus is slightly narrower than the lower isthmus, although the complexity is roughly similar. The size of the lumen in the generated cross sections and the histology images is similar, and the complexity of branching is roughly similar. The cross section shown in C14 is longitudinal, so appears different to the other cross sections.



*Figure 5.16:* Comparison between three generated cross sections A, B & C and three histology images from the corresponding position in the upper isthmus of the oviduct. One image from histology set A slice 16, one image from histology set B slice 79 and one image from histology set C slice 14 are shown for comparison.

Figure 5.17 shows a comparison between the generated cross sections of the ampullar-isthmic junction and the histology images. The complexity of the internal structure is more complex than the upper and lower isthmus, and the size is larger. When compared to the histology images, the branching complexity is similar, with the same limitations as the sperm reservoir, the branches in the histology images are slightly longer, and the lumen space in the middle of the generated cross sections is larger.

Figure 5.18 shows a comparison between the generated cross sections of the upper isthmus and the histology images. The histology images show a significant difference from the generated cross sections. The outer and inner walls of the generated cross sections are far wider than the histology images, and there is a large space within the lumen. This difference is due to the modifications made to the construction algorithm to compensate for the swelling that the ampulla undergoes at the time of ovulation, as described in Section 5.2. This swelling was not present in the histology images. Therefore, a comparison between the cross sections and the histology images is not possible for the site of fertilisation. When compared to the descriptions in literature, Nilsson and Reinius (1969) noted that a normal ampulla was approximately  $300\mu\text{m}$  diameter,

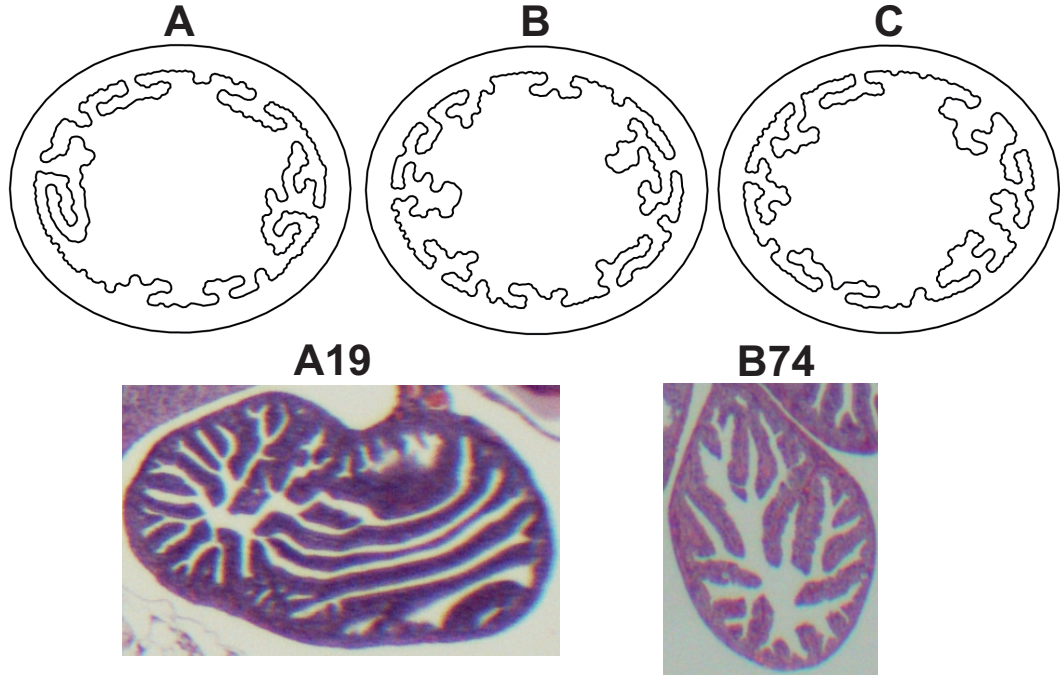


*Figure 5.17:* Comparison between three generated cross sections A, B & C and three histology images from the corresponding position in the ampulla-isthmic junction of the oviduct. One image from histology set A slice 10 and two images from histology set B slices 67 and 68 are shown for comparison.

had many folds and a narrow lumen. This description is consistent with the histology images. Nilsson and Reinius (1969) also noted that when distended by fluid and eggs, the outer diameter was approximately  $500\mu m$  in diameter, had small longitudinal folds and a wide lumen. This description is consistent with the generated cross sections.

Figure 5.19 shows a comparison between the generated cross sections of the pre-ampulla and the histology images. Only image B64 is a true cross section, the remaining images are longitudinal. When compared to image B64, the generated cross sections have similar complexity and branch length. However, the generated cross sections have a slightly larger space inside the lumen.

This type of visual comparison does not provide a quantitative validation of the model, it only demonstrates that the structural features of the model are within the same scale and have similar complexity to the source tissue. More in-depth quantitative validation could be performed. One technique would be to count the number of primary and secondary branches and use this to test for significant differences. This comes with associated issues, such as how to differentiate between a branch and a "bump" in the tissue wall. It would also ideally require a larger number of comparison images. Image based comparison techniques could also be used. Edge detection could be used to identify the lumen outline of the histology image. From this, various metrics could be calculated, such as the lumen surface area, perimeter length and branch thickness variations. This would provide a more quantitative validation of the 2D cross sections



*Figure 5.18:* Comparison between three generated cross sections A, B & C and two histology images from the corresponding position in the site of fertilisation in the ampulla of the oviduct. One image from histology set A slice 19 and one image from histology set B slice 74 are shown for comparison.

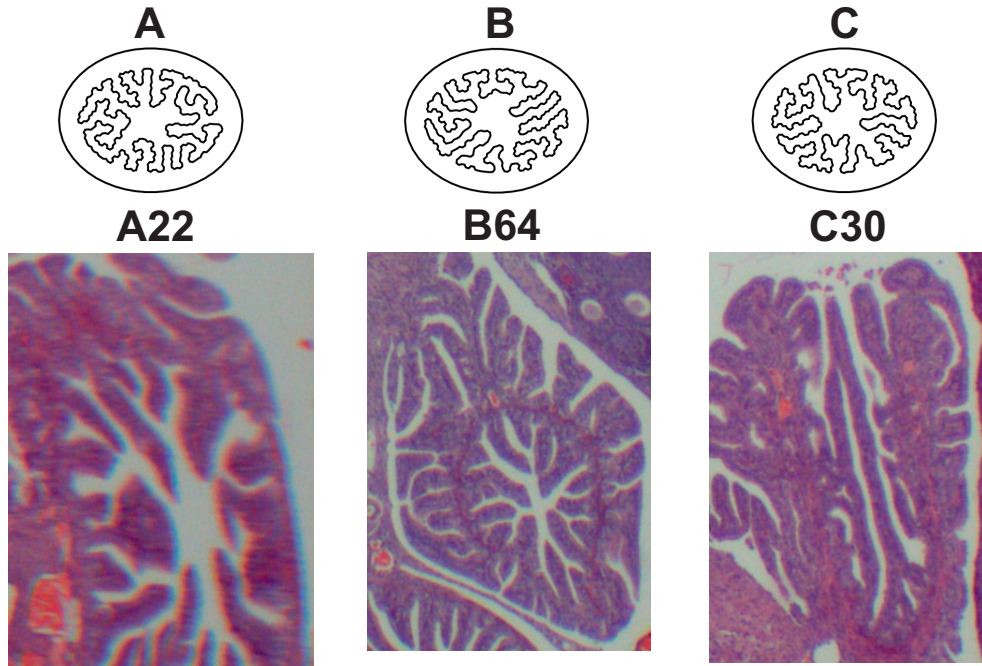
with respect to the histology images. As the aim of this research is to investigate sperm behaviour, and not to generate an exact replication of the oviduct, the visual comparison was deemed sufficient.

### 5.6.2 Lumen path identification

Validation of the path is only possible with direct comparison to the source images. The cluster intensity for each image identified in Section 5.4.1 was visualised in 3D using the same technique defined in Section 5.2.2. This provides a good visual indication of the path that the oviductal tube follows. Three different source data-sets were processed to validate the approach. Figure 5.20 shows all three curves superimposed on the cluster intensity maps. Although not clear from the images due to the complex 3D nature of the curve, all three data sets showed a good fit to the source data.

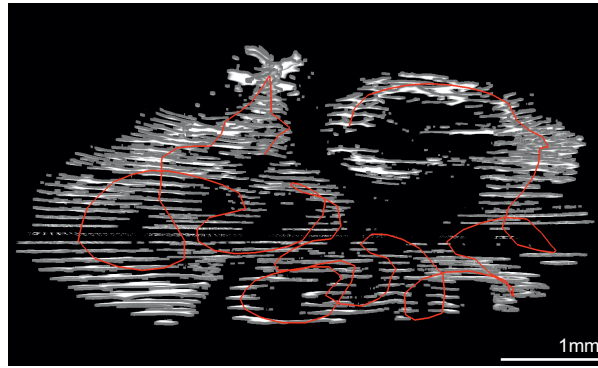
### 5.6.3 Limitations

The 3D models which can be generated using these techniques are reasonably realistic, and provide sufficient detail to simulate the influence that the complexity of the

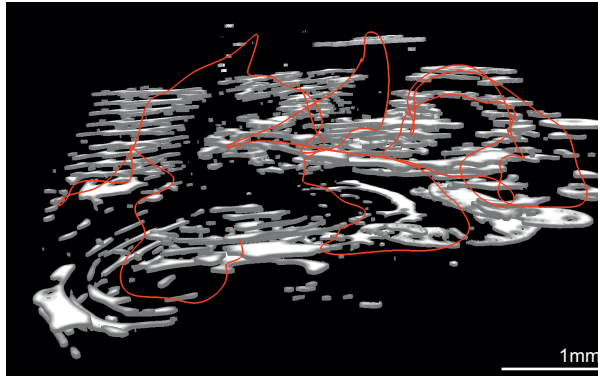


*Figure 5.19:* Comparison between three generated cross sections A, B & C and three histology images from the corresponding position in the pre-ampulla of the oviduct. One image from histology set A slice 22, one image from histology set B slice 64 and one image from histology set C slice 30 are shown for comparison.

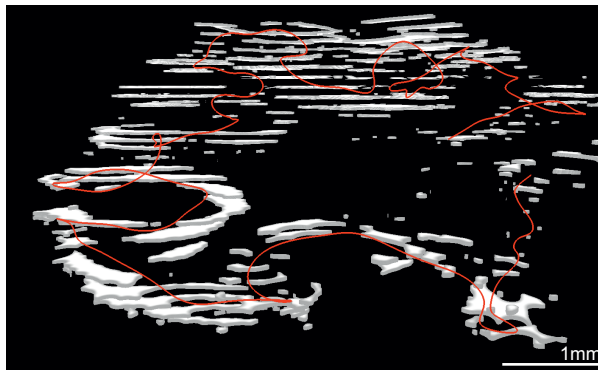
3D environment has on sperm navigation. However, the 3D models also have several limitations from a biological perspective. When examined using Scanning Electron Microscopy (SEM), the fine surface of the oviduct has crypts and branching structures which would influence the interactions of spermatozoa (Yaniz et al., 2006). The surface also contains ciliated cells, the beating of which may also influence the interactions of spermatozoa and secretory cells, which change in size at different stages of the oestrous cycle (Lyons et al., 2006). None of these details are represented in the 3D model. Biological tissue undergoes uneven shrinkage during the histology process, which means that the images used to create the model will not be a true representation of the 3D structure (Beck and Boots, 1972). Although the influence of this on the identified 3D path followed by the tube is likely to be minimal, the influence on the internal folds of the oviduct will be far greater. In reality, the lumen may be wider or narrower in parts, or may have a different structure. The significance of these differences can only truly be investigated once the technology for obtaining small scale medical imaging data has improved. However, it can be partly investigated by modifying the 3D structure to reduce the distance between the internal folds, thus reducing the swimming space within the lumen. This could provide some insight into the impact that histology shrinkage



(a)



(b)



(c)

*Figure 5.20:* Curves identified from 3 data sets superimposed on their intensity maps. Although not visually apparent from the images, the curves match well to the original image data.

has on the model.

Another potential limitation could present itself when sperm agents interact with the surface of the model. Although the model is highly detailed, at its lowest level it is a collection of triangles with edges. When sperm interact with the triangle, the angle of the plane that the triangle falls on will not be a truly accurate representation of the curvature of the real tissue. As sperm to oviduct interactions are based on random reflection, it is likely that the observed effect would be minimal. The effect of this could potentially be reduced by applying a sub-division algorithm to the model. This would increase the number of triangles, and reduce the angle between adjacent triangles. Another alternative would be to estimate the curvature at a particular point in the model based on the normal angle of adjacent triangles and vertices.

### 5.7 3D models for agent based simulations

Having a single 3D model of the oviduct for the agent based simulation could lead to a bias in the results, as an artefact in the model could have a large impact. One way of reducing the potential impact is to use multiple models and then average the results. As three sets of images were used, three different curve paths were identified. Although the internal tissue measurements were averaged across all three samples, each cross section generation run resulted in a unique set of cross sections due to the stochastic natures of the process. The three curves (1, 2 & 3) were cross combined with three sets of cross sections (A, B & C), resulting in nine unique models.

Two additional sets of models were created in order to test the limitations of the approach. The first additional set of models was to investigate the effects of histology shrinkage on the models. The histology process results in uneven tissue shrinkage by approximately 10%. The impact on the oviduct is an increase in the luminal space. In order to compensate for this shrinkage, the generated cross sections were modified by scaling the position of all points by 0.9 and 0.8, and nine models were created for each scaled down set of cross sections. Although not an accurate reversal of what happens during the histology process, it does result in a reduction in swimming space within the lumen, which is the factor of interest.

The purpose of the second additional set of models was to investigate the influence of external bends and internal folds on the simulation in isolation. An additional set of cross sections was generated using the internal and external diameter measurements only, without the internal folds. This was used to generate further set of models. Three models with the external bends but no internal folds, three models with internal folds but no external bends and one model with no internal folds and no external bends

were then created. Full details about the different models constructed for use in the simulations are shown in Appendix C.

### 5.7.1 Oviductal regions

In order to investigate the distribution of sperm within the oviduct over time, the different regions of the oviduct within the 3D model were identified. Figure 5.21 shows how the regions are grouped for analysis, and how they relate to the individual cross sections.

## 5.8 Optimisation

This section describes the performance results for both the image processing and particle physics based simulation. All performance comparisons were carried out using the configuration shown in Table 5.2.

Hardware Configuration
Windows 7 x64
Intel Core 2 Quad Q6700 @2.66Ghz
4GB RAM
NVIDIA GeForce GTX 280 [1GB]

Table 5.2: Hardware configuration used for all simulations.

### 5.8.1 Image processing

A set of 33 images was processed to test the performance difference between the Single CPU and Multi CPU approaches. Each image was processed on a single thread on the CPU, allowing multiple images to be processed at the same time. The process was run using 1 thread, 2 threads, 4 threads, 6 threads and 8 threads, and the results are shown in Table 5.3. Using a single thread, the GPU remains idle while the CPU performs part of the processing. When 4 threads are used, with 1 thread running on each of the 4 cores of the CPU, the GPU can be used while the other threads are performing CPU processing. This results in a performance increase of 1.99 times, with the limiting factor being the GPU. If 6 threads are used, 2 threads can still perform CPU operations while the other 4 threads are waiting for the GPU to return, ensuring that the GPU is constantly in use. When the number of threads used exceeds the number of CPU cores, the performance increase becomes less significant, and using significantly more threads than CPU cores, as shown when 8 threads are used, there is a negative impact on performance. Using this technique, if multiple GPUs were available, the performance

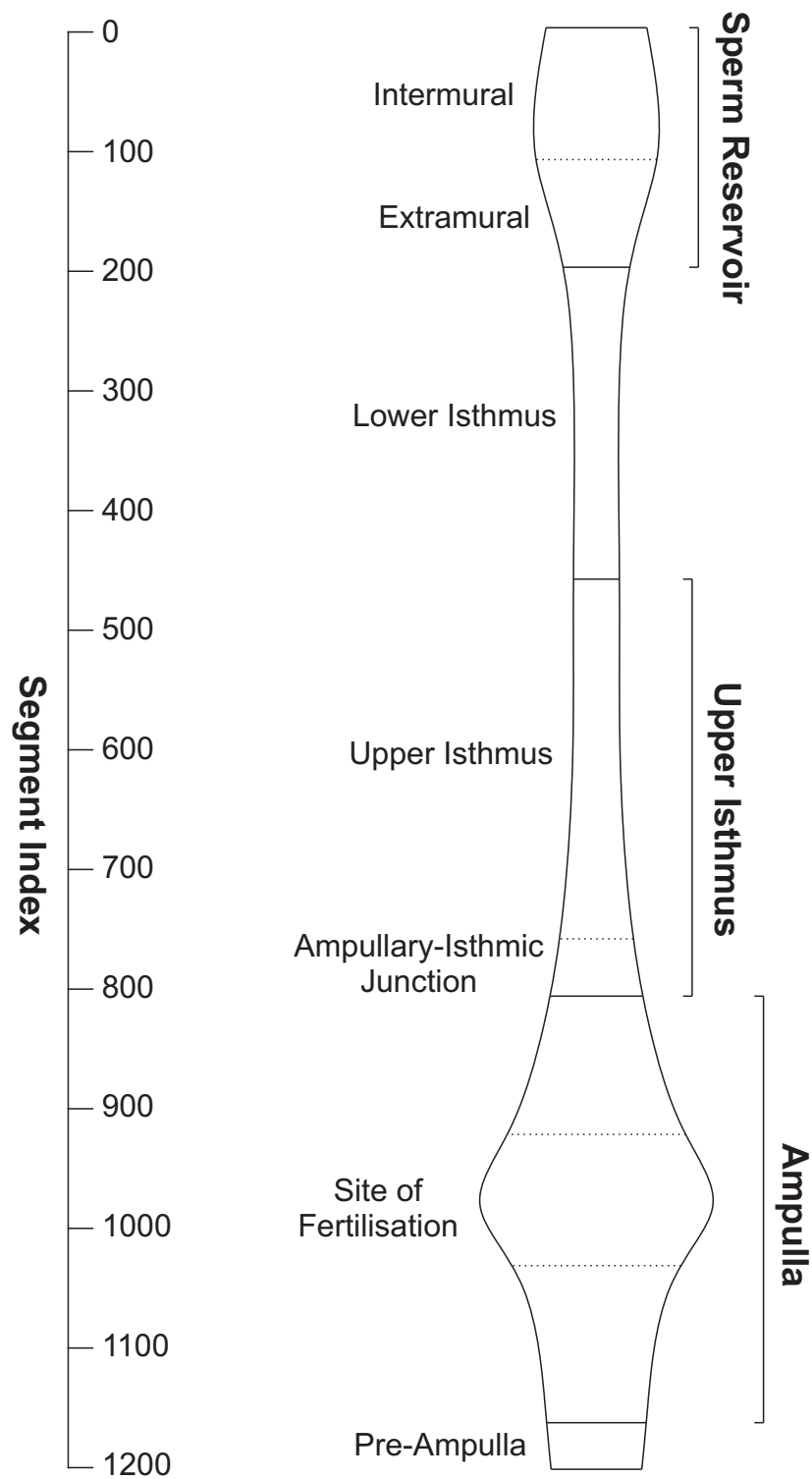


Figure 5.21: The different regions of the oviduct and how they relate to the generated cross-sections. The bold groupings indicate where the individual regions are combined for analysis.



increase is likely to be significantly higher.

No of Threads	Run Time (s)	Performance Increase
1	39.578	x1.00
2	26.032	x1.52
4	19.906	x1.99
6	15.781	x2.51
8	18.594	x2.13

*Table 5.3:* Time taken to complete processing of 33 images with dimensions 1481 by 1304 using the different numbers of threads.

### 5.8.2 Cross-section resolution

The CPU versions of the particle physics based simulation both ran for 9313 iterations and the GPU versions ran for 7619 iterations before reaching equilibrium. This difference is likely due to the way floating point numbers are processed on the different architectures (Kirk and Hwu, 2010). Table 5.4 shows the average time taken to complete each simulation. The GPU times have been scaled up based on the average time for a single iteration to show how long the simulation would have taken to complete if it ran for the same number of iterations as the CPU versions. The Single CPU version took almost 5 hours to reach equilibrium. The Multi CPU version was 3.4 times faster, completing in 1 hour 26 mins, the scaled serial GPU version completed in 7.5 minutes and the vector based GPU version was the fastest, completing in 7.1 minutes, 41 times faster than the Single CPU version. The difference between GPU versions is only slight, and the vector based implementation is only 1.05 times faster than the serial version, however it requires 25% more memory than the serial version.

The use of the GPU to perform the cross-sectional intersection resolution operation gives benefits other than just speed. The GPU is highly optimised for processing large numbers of calculations concurrently. This means that potentially, multiple models could be resolved at the same time without a significant increase in simulation time. This depends greatly on the number of particles within each simulation. Oviducts from other mammals have different lengths and internal complexity. This technique means that for models with a far larger number of cross sections, collisions can be resolved within a reasonable time. A large increase in the number of particles will eventually cause a negative impact on performance, however for this type of simulation, the primary limitation on the GPU is the amount of memory available. When compared with the CPU, the performance will still be significantly greater when a larger number

of particles are used.

Implementation	Iterations	Run Time (hh:mm:ss)	Performance Increase
Single CPU	9313	04:54:33	x1.0
Quad CPU	9313	01:26:22	x3.4
GPU (serial)	7619	00:06:11	x47.9
GPU (serial) Scaled	9313	00:07:33	x39.2
GPU (vector) Scaled	9313	00:07:08	x41.4

Table 5.4: Relative number of iterations and time taken for the particle physics simulation to reach equilibrium for each implementation.

## 5.9 Conclusion

The main aim of this chapter was to generate a set of unique 3D computational models of oviducts based on histology images. The stacked images were visualised using a simple technique to give the illusion of volume. Each image was then processed, using both multi-core CPU and GPU image processing techniques, to identify and extract cross-sections of the oviductal lumen. A set of elliptical regions encapsulating the cross-sectional lumen were identified and then connected using a set of rules. Domain specific knowledge of how the oviductal tube is formed is used to simplify the rules for creating the complete 3D path. Measurements taken manually from the original images were related to a point along the 3D curve and used to create a series of linear 2D cross-sections. These cross-sections were bound to the 3D path and cross-sectional intersections and tubular collisions were automatically resolved using a novel particle physics based technique running on single-core CPU, multi-core CPU and the GPU. The relative performance of these techniques was then compared. The advantages of using multi-core CPU and GPU high performance computing techniques to perform a computationally intensive tasks for the complete construction of a 3D environment suitable for computational modelling has also been demonstrated.

Although models generated using these techniques are not as accurate as a more exact reconstruction method, there are several advantages. Simulations can be run using multiple unique models with similar properties, which means that the effect of individual variations on the simulation results can be mitigated. Different sets of models can be created with the same external structure but with a different internal structure. The complexity of the internal structure can be varied, making it more complex or less complex than normal. These models can then be used to investigate the influence that the internal complexity has on the movement of sperm within. Models

with different external structures but similar internal structures can also be created, to test the influence that bends have on internal interactions. As the models are based on measurements and features extracted from histology slides, they are grounded in reality. When compared visually to histology images, the generated cross sections have similar size, complexity and internal structure. Although differences are observable, the generated models are sufficiently accurate to investigate the influence of the 3D structure on sperm behaviour.

Chapter 6 describes the implementation of the agent based model of sperm movement within the generated 3D oviduct models.

## Chapter 6

# Computational model implementation

This chapter describes the process of creating an agent based model of sperm movement within the female reproductive tract. First, an analysis of the available ABM frameworks was performed to determine if an existing framework was suitable for this type of model. Next the implementation of the agent model of sperm behaviour is described and the technical aspects of model construction are presented. Finally the performance benefits of the optimisation techniques applied to the model are presented.

### 6.1 Selecting an ABM framework

When considering how to implement an agent based model, there are two main options, to write a custom system or to use an existing agent modelling framework. Writing a custom system allows complete flexibility. The choice of programming language, techniques and technologies used can all be tailored to the problem being addressed. A custom implementation could also take full advantage of modern hardware, such as multi-core programming frameworks like OpenMP (Dagum and Menon, 1998), or even make use of the parallel processing power of modern graphics cards using CUDA (NVIDIA, 2007) or OpenCL (Khronos, 2009). However, there are also significant disadvantages to writing a custom system. The main disadvantage is the amount of time required for development. The overall framework would need to be designed, developed and comprehensively tested to ensure that it functioned correctly. The resulting system would still be unproven, and could include bugs which may impact the results from any simulation run. Although writing a custom system is the most flexible, it is not feasible to spend a large amount of time on this task when several existing frameworks provide an adequate set of features.

A large number of agent based frameworks are available, many of which are written to solve a particular class of problem. An evaluation of the available tools and technolo-

gies for performing agent based modelling was carried out with the aim of finding an appropriate framework for creating the agent based model of sperm behaviour. Many frameworks attempt to provide a complete set of tools for creating a model, analysing the results and visualising the interactions. However, it is more useful to separate these and choose tools dedicated to each task. This being the case, the main criteria for the evaluation will be the modelling capabilities, performance and the ability to generate data in a format suitable for external analysis and visualisation.

The majority of agent based frameworks provide a set of common tools which are used to define agents and control their interactions through scheduled actions. Typically, each agent is implemented as a class, and the properties of the class are the agent properties, the functions of the class are the agent functions. The order and timing of agent function calls is explicitly specified in the initialisation code, and lists of agents are explicitly managed. Agents have direct access to each other, and can read each others properties at any moment in time.

### 6.1.1 Framework summary

Railsback et al. (2006) provide an evaluation of five of the more popular agent based frameworks, NetLogo (Wilensky, 1999), MASON (Luke et al., 2004), Repast (ROAD, 2005), Swarm (Objective-C) and Swarm (Java) (SWARM, 2009). Sixteen versions of a simple model called StupidModel (Steve Railsback and Grimm, 2005) were implemented in each framework. The different versions of the model progressively increase in complexity, testing various aspects and capabilities of the different frameworks. Based on this evaluation, NetLogo was excluded due to its simplistic nature and the Objective-C version of Swarm was excluded due to its increased complexity compared to the Java equivalent (Railsback et al., 2006). Two additional frameworks, Flexible Large-scale Agent Modelling Environment (FLAME) (Coakley, 2007) and a GPU based version, called FLAME GPU (Richmond et al., 2009) were evaluated using the StupidModel to give a general comparison with the more mainstream frameworks. Below is a brief summary of the different frameworks.

#### 6.1.1.1 Swarm

The Java version of Swarm is a Java wrapper for the original version of Swarm, written in Objective-C. Designed for domain independence, Swarm allows a hierarchy of models to be implemented, thus simplifying process of multi-scale modelling by allowing interactions between models with different spatial and temporal scales. A basic model in Swarm is constructed with a top-level swarm, and lower level swarms defining more

complex interactions. Swarm contains built in visualisation and data analysis tools, allowing real time interaction with the simulation. Agent functions are scheduled as actions to occur at specific intervals, and are registered by specifying the class type and the function name as a text string.

#### 6.1.1.2 Repast

Repast was originally developed to be a pure Java implementation of Swarm, but has since become an independent framework in its own right. It contains a complete library of analysis and visualisation tools, however it is targeted at social sciences, and many of the tools are domain specific. Like Swarm, agent functions are scheduled as actions to occur at specific intervals using the same registration method.

#### 6.1.1.3 MASON

MASON is a pure Java implemented framework, which is designed to be fast. It also contains a set of analysis and visualisation tools. Scheduling actions is slightly more complex in MASON than the other Java frameworks. To implement a function, a class must implement an interface, and classes implementing the interface can be registered with the scheduler. MASON also allows agent functions to be parallelised, dividing the workload between multiple CPU cores.

#### 6.1.1.4 FLAME

FLAME, which stands for Flexible Large-scale Agent Modelling Environment, is a C based modelling framework designed to take advantage of massive parallel processing power of grid computers (Coakley, 2007).

Agent modelling frameworks commonly process lists of agents in serial, one after another. FLAME is different in that agents are processed concurrently on different CPU cores. This introduces a couple of problems with the modelling process. If agents are processed serially, then the state of all agents is accurate, and agents can interact directly with each other without any problems. However, if agents are being processed concurrently, the state of multiple agents could be changing at the same time. Therefore when one agent is being updated, the state of other agents cannot be accurately determined.

In order to counter this limitation, FLAME use a messaging architecture. Instead of agents directly interacting with each other, they generate a set of messages, which are stored in a list called a message board. Once all messages have been generated, the

message board is locked and a set of agents marked as consumers of the message are able to iterate through all messages in the message board, and are able to process any messages they are interested in. Each agent is responsible for updating their own memory, thus removing the issues surrounding deadlock and multiple agents accessing the same memory location.

FLAME is a template generation framework, and uses procedural programming instead of object oriented programming. It is based on X-machine theory, and therefore only consists of agents, states and transitions (Coakley et al., 2006). FLAME has a fixed structure, and the user only has to specify code related to the model. Code for managing state transitions, agent lists, memory allocation, data persistence and kernel execution are all automatically generated from the XML template. This greatly simplifies model development, allowing the model developer to focus solely on the model without worrying about the complexity of parallel programming.

To create a model in FLAME, a model description is created using an XML file to define agents, their properties, the functions that act on those agents and messages that pass between agent types. A separate file written in C provides the implementation of agent functions. The XML file is then parsed before compilation, and a set of source code files based on the agent definitions is generated. This generated source code contains all the ‘boiler plate’ code which manages agent arrays, type definitions, interactions and scheduling. These source files are then compiled to generate the final model.

When running a model, an XML initialisation file is passed in as a parameter. This file determines the number of agents to run within a simulation and the starting values for each agent. Every iteration, a sequentially numbered XML file is generated with the current state of all agents. Unlike the other agent based frameworks, FLAME does not provide any visual interface for running and interacting with the model.

#### 6.1.1.5 FLAME GPU

FLAME GPU is a re-engineered version of FLAME which runs entirely on the Graphics Processing Unit (GPU) using CUDA (Richmond and Romano, 2011). This allows FLAME models to be run using relatively inexpensive consumer NVIDIA graphics cards (Kirk and Hwu, 2010). Graphics cards (GPUs) are primarily designed for displaying 3D images on the screen at a high frame rate. They are highly optimised for performing a large number of fast mathematical calculations in parallel. In recent years, the GPU has been used for scientific computing, as the fast, parallel nature of the hardware is ideal for performing certain types of computational simulation (Kirk and Hwu, 2010). CUDA is a set of extensions to the C programming language which

makes it simpler to access the memory and run functions, which are called kernels, on the GPU.

The model syntax of FLAME GPU is very similar to FLAME, an XML file is used to define the agents and environmental variables within the model, and CUDA code is used to implement the functions. CUDA is similar to C, with a few graphics card specific extensions. Similar to FLAME, FLAME GPU also has a fixed structure and a similar method for defining the model.

To create a model using FLAME GPU, an XML based definition file is created which describes each agent and their properties, how agents interact with each other, and provides a set of function declarations. The implementation of each function is then written in CUDA, with each function forming a CUDA kernel.

### 6.1.2 StupidModel

To provide comprehensive comparison between the five potential frameworks, several versions (1 - 16) of the StupidModel (Steve Railsback and Grimm, 2005) were implemented in both FLAME and FLAME GPU. As FLAME does not have visual display tools, and the evaluation criteria focuses on model development only, the visual and interactive additions to the stupid model have not been implemented. This means that version 4, which introduces inspection of values mid-simulation and versions 6 and 13, which requires the generation of graphs, were not used for evaluation. The comparison between FLAME / FLAME GPU and the Java based frameworks is based on the descriptions in Railsback et al. (2006).

#### 6.1.2.1 Version 1: mobile agents

This is a simple model with a single agent type that represents bugs on a 2D grid. The bugs move randomly and two bugs cannot reside on the same location. Each location on the grid can contain a bug, and each iteration, all bugs are moved to a free space sequentially, ensuring that two bugs do not move to the same location.

Repast, Swarm and MASON implement the environment as a 2D grid, which is a built in class. As two bugs cannot reside on the same location, collisions are handled by inspecting the current location of other bugs, and only moving to unoccupied locations. FLAME / FLAME GPU does not have the concept of a grid, only agents with position coordinates. The parallel nature of FLAME / FLAME GPU means that handling discrete collisions is significantly complex, requiring either collisions to be resolved linearly, or iteratively, with the potential for an infinite loop to be encountered. This



highlights a potential limitation of the FLAME / FLAME GPU framework which needs to be taken into consideration when the model is created. Instead, this version of the model allows multiple bugs to reside on the same grid location.

#### **6.1.2.2 Version 2: agent growth**

A second action was added to increase the size of a bug agent every iteration. This was trivial in all frameworks with the exception of MASON. As MASON required a specific interface to be implemented for an action, only one function can be implemented per class. For complex models with a large number of functions, this could make the code difficult to manage. The more recent versions of Java allow anonymous classes to be defined from interfaces, which means that this is less of a problem.

#### **6.1.2.3 Version 3: habitat cells and resources**

A set of agents representing the local habitat at each location in the grid was added to the model.

In Repast, Swarm and MASON the model was modified to assign a single cell to each location in the grid, and then allocate a bug to the habitat cell, instead of the grid location. Bug growth is based on consumption of resources from the habitat cell, with bugs directly accessing the cell it is residing on to query the amount of food available.

As FLAME / FLAME GPU does not have the concept of a grid, each habitat cell was created as a new agent and interactions between agent types were based on messages passed between them. This resulted in a significant impact in performance, due to the introduction of 100,000 new agents, each with multiple associated messages. The limitation of multiple bugs being able to inhabit the same location results in a potential negative value for resources in a habitat.

#### **6.1.2.4 Version 5: parameters and parameter displays**

The ability to view and modify parameters before and during the simulation is investigated. Repast, Swarm and MASON allow any parameter to be specified before the model is run and then dynamically altered during execution.

FLAME / FLAME GPU allows parameters to be defined prior to model execution, either as a static constant in the main functions file, or as a parameter in the XML definition file. Dynamic modification of parameters during model execution is not permitted. The inability to change the values of variables mid-simulation is not necessarily

a limitation however, as doing so introduces the possibility of human error.

#### **6.1.2.5 Version 7: stopping the model**

The ability to define conditions under which the model will stop running was investigated.

Repast, Swarm and MASON allow the model execution to be interrupted when specific conditions are met.

FLAME / FLAME GPU does not allow the model executed to be interrupted within code. Instead, the number of iterations that the simulation will run for is specified when the model is executed. As the state of each agent is written to an XML file, if a simulation does not run for enough iterations, the final generated output file can be used as an input in a new simulation run, thus effectively continuing the simulation from the last stop point.

#### **6.1.2.6 Version 8: file output**

The ability to write the data generated to a file is investigated. Repast has built in file writing capabilities, but Swarm and MASON require custom output code to be written, however this is simple due to the java.io packages.

FLAME automatically generates XML files each iteration containing the state of each agent.

#### **6.1.2.7 Version 9: randomisation of bug movement**

The ability to randomise the order in which bugs are processed is implemented. Both Repast and Swarm allow randomisation of agents. FLAME / FLAME GPU does not randomise the order of agents. However, as FLAME / FLAME GPU run agents in parallel, they are processed at the same time, not in any specific order. For simulations with a large number of agents, groups of agents will effectively be processed at the same time. This means that the agents at the beginning of the list will always be processed before the agents at the end of the list. However, as interactions between agents are performed using messages, the order that agents are processed will not have an impact on the simulation.

#### 6.1.2.8 Version 10: size-ordering of bug movement

The ability to sort the list of agents based on a particular property values was investigated. Repast, Swarm and MASON allow sorting of agents. FLAME / FLAME GPU does not allow sorting of agents, however as mentioned above this is not a limitation due to the way interactions between agents are handled.

#### 6.1.2.9 Version 11: optimal movement

The random movement of bugs was modified to inspect neighbouring cells and choose the cell with the most resources. Repast and MASON provide a method for automatically identifying neighbouring cells, Swarm and FLAME / FLAME GPU required custom code to be written. In all implementations, neighbouring cells were iterated over to find the cell with the most resources. Where multiple cells had the most resources, the chosen cell was randomly selected.

#### 6.1.2.10 Version 12: mortality and reproduction

The ability to add and remove agents part way through the simulation was introduced. When a bug reached a specific size, it produced a single offspring and then died.

In Repast and Swarm, it was not possible to change the agent list during a scheduled function. A separate list of new agents was maintained, and the master list was updated at the end of an iteration. In MASON, the agent list was copied and the actions scheduled to run on the copied list. The master list was then changed based on the outcome of each agent interaction.

In FLAME / FLAME GPU, this process is far simpler. The automatically generated *add new agent* method is called to add agents. To remove an agent, a function running on that agent returns 0, and it is automatically removed from the internal list of agents. In FLAME GPU, there is a limitation that only a single new agent can be created per agent function.

FLAME / FLAME GPU have a problem with assigning unique identifiers to new agents. As the code is running in parallel, the current number of agents is not known, so determining the next free ID is not a simple process. One solution is to have an environment agent with a single instance. This agent is responsible for assigning unique identifiers to agents in response to messages.

#### 6.1.2.11 Version 14: random normal initial size

The ability to populate agent values from specific statistical distributions was investigated. Both Repast and Swarm contain this functionality, but MASON, FLAME & FLAME GPU do not. As the initial value for agents in FLAME / FLAME GPU is specified externally in an initial XML file, the inclusion of this functionality within the framework would not be beneficial. As the XML files can be automatically generated by any programming language, the initial value of agents could easily be specified from a specific statistical distribution.

#### 6.1.2.12 Version 15: habitat data read from a file

The ability to import a text file containing the environmental dimensions and resources availability for each habitat location was investigated. Repast, Swarm and MASON required custom code to be written to import the file. Importing the file directly was not built into FLAME, although it could be implemented in FLAME GPU, as it contains a pre-simulation initialisation function. Instead the text file was used to populate the habitat descriptions in the initial XML file.

#### 6.1.2.13 Version 16: predators

The model was modified to include a new type of agent, a predator, which moves randomly and consumes a bug if it moves to the same location. Repast, Swarm and MASON required changes to the underlying model to accommodate multiple agents at a single habitat. FLAME / FLAME GPU simply required a new agent type to be added, and the specification of corresponding message functions for the Predators and Bugs.

#### 6.1.2.14 Discussion

Using the StupidModel as a target for evaluation has highlighted several limitations with both the FLAME and FLAME GPU when compared to the Java frameworks. Inter-agent collision detection is non-trivial in FLAME / FLAME GPU, and there are no built in structure to manage commonly used environments such as 2D grids. Simulations in FLAME / FLAME GPU cannot be interrupted once a specific set of conditions are met, and simulations must be run for a specific period of time. In FLAME / FLAME GPU adding and removing agents is trivial, however, assigning unique identifiers to agents is non-trivial due to the parallel nature of the frameworks. In FLAME / FLAME GPU, creating the model and adding new agent types is very

simple as the code for managing lists of agents and their interactions is automatically generated.

Despite the limitations of FLAME and FLAME GPU which were highlighted by the comparison, both frameworks have several advantages that the StupidModel has failed to highlight. When creating a model in FLAME, the user focuses entirely on the model. The framework automatically handles everything else. This means that it is very easy to learn and use, and reduces the scope for errors. FLAME models can also be run in parallel. FLAME allows the model to be run over a distributed network or High Performance Grid system, and FLAME GPU uses the parallel nature GPU processors to run the model in parallel on a single computer. This can potentially result in significant performance gains. Further evaluation was needed to determine which framework was the most viable.

### 6.1.3 Basic sperm model

Although the StupidModel implementations gave a good insight into the capabilities and limitations of each framework, the type of model implemented is very different to the model required for simulating 3D sperm motion. A simple model of sperm motion within a 3D environment was devised and implemented in all five frameworks.

#### 6.1.3.1 Evaluation criteria

Three different criterium were used to evaluate the simulation. The agent model of sperm behaviour is quite complex, with several different processes and agent states. Therefore, the agent framework should make implementing the logic of the model as simple as possible, while reducing the potential for errors. As identified in Chapter 4, each simulation step represents 1 second, and the simulated time period will be several hours. Therefore, a large number of iterations are needed. As the important factor in the simulation is the interaction with the 3D environment, the selected framework needs to be able to perform complex mathematical calculations efficiently. Due to the large amount of variability in the system, each set of simulation conditions needs to be replicated multiple times. Therefore, the total simulation time needs to be low. As there are a large number of iterations, the framework needs to be able to manage the massive amount of data generated. These criteria are outlined below:

**Ease of implementation** How easy it is to implement the logic of the model without significant implementation of the ‘boiler plate’ code to manage agents interactions?

**Fast simulation time** Can the simulation framework handle large numbers of agents and how fast does each simulation step take?

**Data management** can the framework generate data in a format that is suitable for analysis?

### 6.1.3.2 Model description

The model contains a set of sperm agents with simple behaviour, moving within and interacting with a 3D environment. The sperm agents have two different states, moving forward and turning. This is a gross simplification of the progressive and non-progressive states described in section 4.3. The sperm agents move forward until they collide with the 3D environment. Then they switch to turning behaviour, turn in a limited, random direction and move forward at half speed. If the agent does not collide with the 3D environment, then they switch to moving forward behaviour, otherwise turning behaviour is maintained.

The implementation of the model was the same for each framework being evaluated, with a few specific differences. FLAME does not provide a pre-processing step, so the 3D environment data was embedded in a pre-compiled static array. Java does not allow large static arrays to be pre-defined, but Swarm, Repast and MASON allow a pre-processing step, so the 3D environment data was loaded from a file. For FLAME GPU, the 3D environment was embedded in a header file. The memory was then copied over onto the GPU card in a pre-processing step and accessed as a texture. The FLAME GPU framework was the most flexible, as the data could also have been loaded from a file in the pre-processing step. Both FLAME and FLAME GPU also had the option of embedding the environment data into the 0.xml startup file, but this would have significantly increased the size of the initial file, and all output files.

The algorithms used within the model were initially implemented in C for use in FLAME, and then ported to CUDA for FLAME GPU and Java for Repast and Swarm to ensure consistency of the implementation. For use with C, a set of vector maths functions were required to perform the calculations. To simplify development, an existing library (Baker, 2006) implemented in C++ was modified for compatibility with pure C. For the FLAME GPU implementation, vector functions provided by the `cutil_math.h` extensions header were used. For the Java implementations, the built in `javax.vecmath` classes and functions from the Java3D package were used.

This does however give the FLAME GPU implementation an advantage, as the vector functions from CUDA are optimised for the hardware, whereas the implementation in

C is not. The same simulation was run with 1,000, 10,000 and 100,000 agents for 100 iterations in all frameworks with 3 replicates of each. The FLAME version was run under several different conditions. It was run in serial mode under Microsoft Windows, and then in parallel mode under Linux on the Iceberg High Performance Grid, which is a University based research facility and part of the Whiterose High Performance Grid. On the Grid the simulation was run using 1, 2, 4, 8, 16 and 32 cores. Table 6.1 shows the hardware configuration for both the Microsoft Windows simulations and the Grid simulations.

<b>Hardware Configuration (Microsoft Windows)</b>
Windows 7 x64
Intel Core 2 Quad Q6700 @2.66Ghz
4GB RAM
NVIDIA GeForce GTX 280 [1GB]
<b>Hardware Configuration (Iceberg Grid)</b>
64-bit Scientific Linux
568 processor cores
435 GFLOPs peak performance

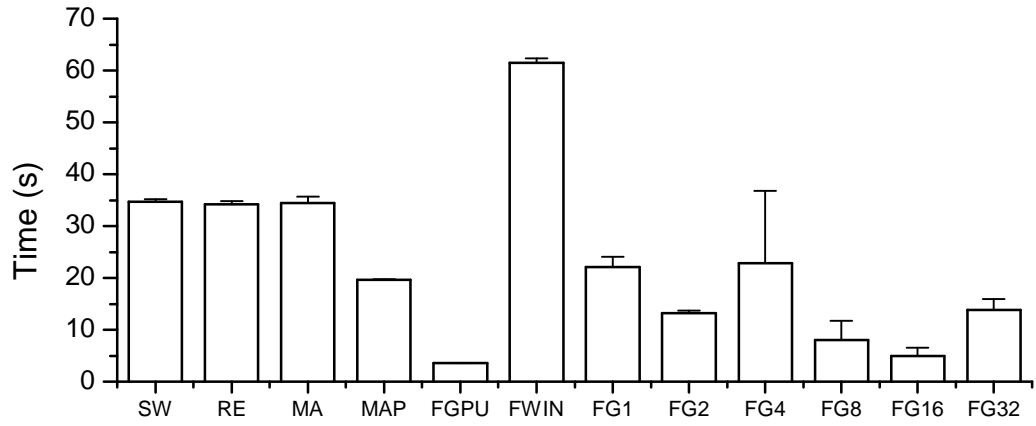
*Table 6.1:* Hardware configuration used for all simulations.

The time each simulation took to complete is shown in Figure 6.1. The fastest 4 implementations, which were FLAME GPU and FLAME running in parallel on the Grid using 8, 16 and 32 cores are focused on in Figure 6.2.

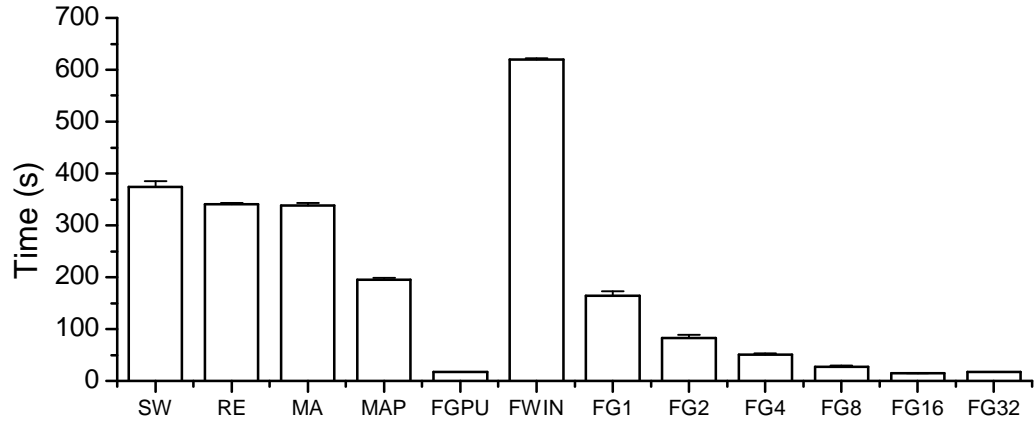
For 1,000 agents, all three serial Java frameworks showed similar performance results. When increased to 10,000 agents. Swarm was slightly slower than Repast and MASON, but all three scaled linearly, so the per agent simulation time was consistent with 1,000 agents. When run for 100,000 agents, both Repast and MASON showed the same linear scalability. However, it appears that Swarm encounters a significant problem when the number of agents becomes very large, with the runtime jumping to almost twice as long as the other Java frameworks.

Running MASON in parallel was around 40% faster than the serial version for all numbers of agents.

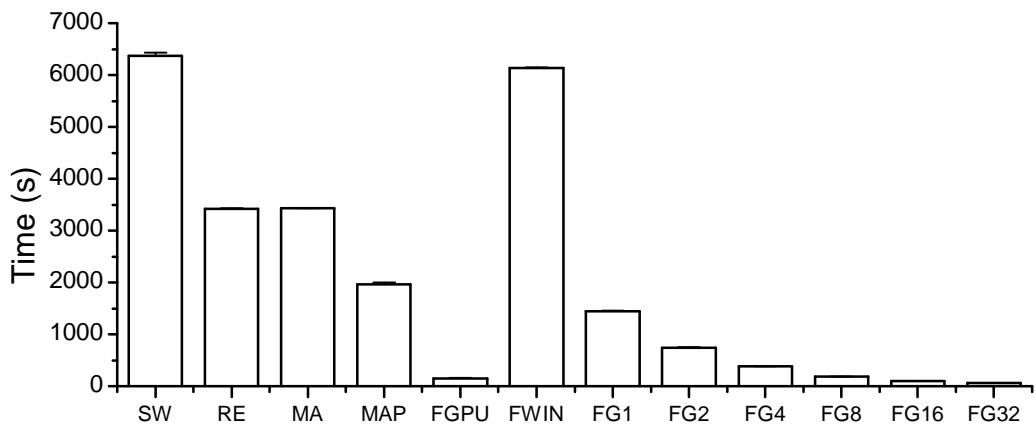
With regards to FLAME, the performance figures shown in Figure 6.1 are a little surprising. As FLAME is written entirely in C, the expectation was that it should always be faster than the Java based frameworks. However, the Java based frameworks were consistently faster than the Windows implementation of FLAME, with the exception of Swarm with 100,000 agents.



(a)



(b)



(c)

Figure 6.1: Relative performance of each framework when run for 100 iterations with (a) 1,000 agents, (b) 10,000 agents and (c) 100,000 agents. SW = Swarm, RE = Repast, MA = MASON, MAP = MASON in Parallel, FGPU = Flame GPU, FWIN = Flame in serial under Windows, FG1 - FG32 = Flame run on the Linux Grid using 1 - 32 cores. Error bars show S.D.



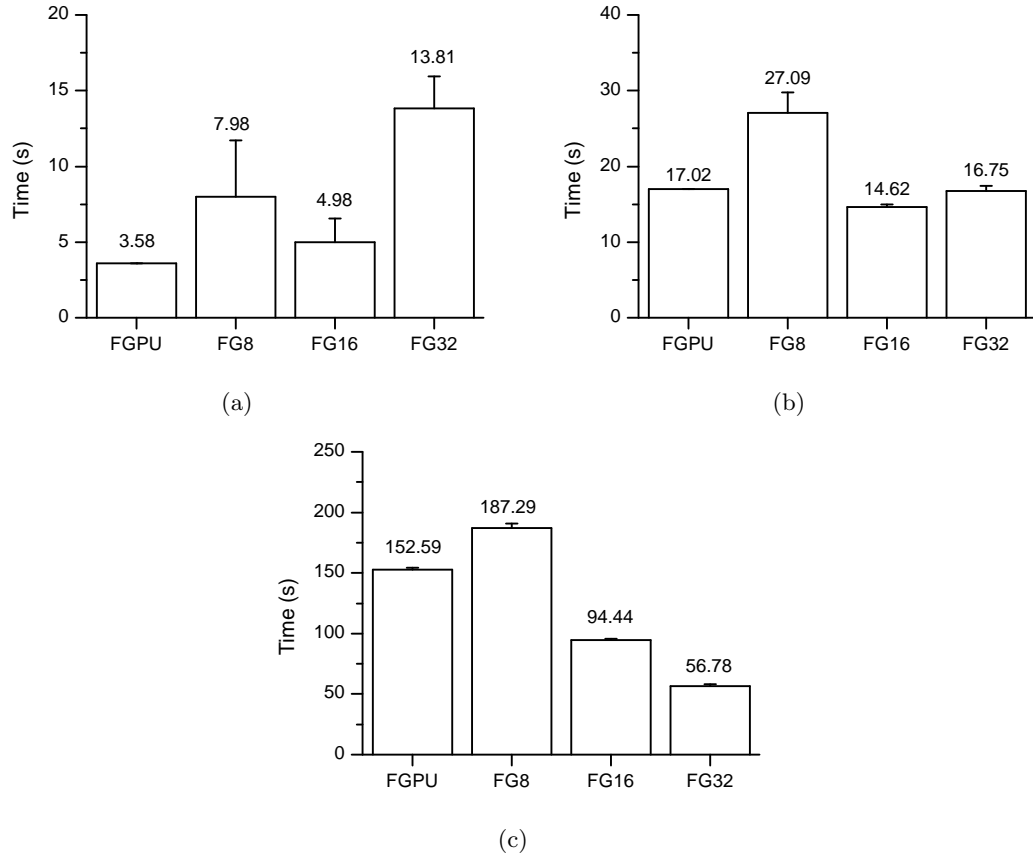


Figure 6.2: Relative performance of the fastest four frameworks when run for 100 iterations with (a) 1,000 agents, (b) 10,000 agents and (c) 100,000 agents. FGPU = Flame GPU, FG8 - FG32 = Flame run on the Linux Grid using 8 - 32 cores. Error bars show S.D.

Running the simulation on the Grid showed mixed results. As FLAME was designed for parallel processing, it was expected that the performance would increase significantly when run on a Grid system. For 1,000 agents, the simulation time was varied for multiple cores. For 4 cores, one of the simulation runs took twice as long to complete as the other two. This could be due to high load on the grid at the time the simulation was run. In general, the simulation run time decreased as the number of cores increased. For 32 cores, the simulation time increased significantly when compared to 16 cores. For 1,000 agents, dividing by 32 cores results in only 31 agents per core. Therefore the communication overhead between cores becomes too large in comparison to the time taken to process the agents.

For 10,000 and 100,000 agents, running the simulation of multiple cores produced on average 45% increase in speed as the number of cores is doubled. For 10,000 agents, the increase from 16 to 32 cores resulted in a slightly increased running time. FLAME GPU

was consistently faster than all the other frameworks with the exception of FLAME on the Grid with 16 and 32 cores for 10,000 and 100,000 agents.

#### 6.1.4 Discussion

Reviewing the evaluation criteria, framework selection was based on (a) how easy it is to implement the model, (b) the framework performance with a large number of agents and (c) ability of the framework to generate data.

In terms of complexity of implementation, FLAME and FLAME GPU were the simplest to implement, as only the model description in XML and functions in C required implementation. All three Java frameworks required classes to be written to manage the list of agents and to perform operations which both FLAME and FLAME GPU handle automatically. FLAME and FLAME GPU did not require any changes to the code to run in parallel, whereas MASON required the scheduling method to be changed, and the agents to be manually divided between the threads. Neither FLAME or FLAME GPU required the ‘boilerplate’ code required to manage lists of agents, schedule events or control interactions. FLAME required custom 3D maths functions to be written, which were built into FLAME GPU and the Java frameworks.

Performance wise, FLAME GPU is superior, even though FLAME on the Grid is faster when a larger number of cores are used. Potentially, the Iceberg Grid has 568 cores available. However, the Grid is shared by many groups in the University and is usually under high load. This means that to run a single simulation, the job gets entered into a queue and will run once it reaches the top. If the Grid is under high usage then it may be several hours before the job reaches the top of the queue. All the other frameworks allow the simulation to be started immediately. Therefore, if the delay is factored into the simulation time, FLAME GPU becomes significantly faster than FLAME.

For data output, the three Java frameworks are superior. FLAME and FLAME GPU generate data in XML format. With a large number of agents running for a long simulation time, this will generate a massive amount of data and as XML files are plane text, each results set will be extremely large. The Java frameworks would be easy to write the data out to a binary format, whereas FLAME and FLAME GPU would require modifications to the underlying framework.

Based on the evaluation criteria, FLAME GPU stands out from the other frameworks due to its ease of implementation and the significant performance increase in relation to the other frameworks. How the limitations in the file output were addressed is discussed in Section 6.4. The massively parallel capabilities of graphics cards and their

optimisation for 3D mathematical calculations made this an ideal choice for this type of model.

## 6.2 Model implementation

This section describes the computational realisation of the conceptual model of sperm behaviour. As described in Chapter 4, the model requires two main components, a set of 3D computational models of the oviductal environment and an agent based model of sperm behaviour.

### 6.2.1 Representation of 3D oviductal geometry

The environment is an accurately scaled 3D model of a mouse oviduct, which appears as a convoluted tube approximately  $17,000\mu m$  in length, and is subdivided into sequential segments, as described in Chapter 5. The final models have over 600,000 triangles, divided into 1,200 sequential slices. The regions of the oviduct (i.e. the isthmus and the ampulla) were identified in relation to the individual slices. This information was used to identify the current regional location of individuals within the environment. All position and distance values are in  $\mu m$  and the 3D environment is scaled with a mapping of 1:1 between  $\mu m$  and 3D units. The environment is limited to the extents of the tube of the oviduct model.

### 6.2.2 Agent parameters

This section describes how the properties for the sperm and oocyte agents, as initially defined in Section 4.3 and listed in Table 4.2, are implemented. Within this section, the properties from Table 6.2 and Table 4.2 are referenced in a **Teletype Font**.

In order to track the movement of agents in 3D space, each agent requires a **[Current Position]** and **[Current Direction]**, as defined in Table 4.2. The **[Current Position]** and **[Current Direction]** of each sperm agent is encoded in a 4x4 transformation matrix (**[T Matrix]** in Table 6.2), which simplifies transformations and calculations in 3D space. A 4x4 transformation matrix is commonly used in 3D graphics to represent the position, rotation and scaling which has been applied to an object in 3D space (Watt, 1999). Operations to transform the position or to apply further rotations to the object can be encoded as a transformation matrix, and then applied to the object matrix using matrix multiplication. As the scale of the agents in the simulation is fixed, the scaling portion of the matrix will always be the same. In order to save agent memory, the matrix is stored as a 3x4 matrix (12 element float array) with the scaling portion

Sperm Parameter	Data Type	Description
ID	integer	Unique agent identifier
T Matrix	float[12]	4x4 transformation (16 element) matrix without scaling so reduced to 12 elements
Activation State	integer	bitwise flags for Alive State, Capacitation State, Movement State and Collision State
Oviduct Segment	integer	the current segment within the environment
Attached Oocyte ID	integer	The ID of the oocyte that this sperm is attached to
Attached Oocyte Time	integer	The iteration time that the sperm made an attachment to the oocyte
Remaining Lifetime	integer	the remaining lifetime for capacitated sperm
Movement State Timer	integer	The remaining time before a change in movement state is made

Oocyte Parameter	Data Type	Description
ID	integer	Unique agent identifier
Current Position	float[3]	x, y, z position in 3D space

Table 6.2: Sperm and oocyte agent parameters.

removed, and re-inflated to the full 16 element array for calculations.

The [Capacitation State], [Movement State], [Collision State] and [Alive State], as defined in Table 4.2, and linked in Figure 4.3, are encoded as a single integer ([Activation State] in Table 6.2). The individual bits of the integer are used as boolean flags, allowing any combination of states to be represented in a single number. The [Alive State] is not explicitly modelled. Instead, the absence of a [Capacitation State] (i.e. all capacitation state bits are zero) is used to indicate a sperm agent with the *dead* state. Figure 6.3 shows how the first few bits of the integer bit sequence are used to store all three states. For example, if bit position 1, 4 and 5 are set to 1, then this will represent the state *alive*, *capacitated*, *non-progressive* and *free swimming*, and will have an integer value of 50. An additional [Capacitation State] value was introduced for the computational model (*post capacitated*), which is used to distinguish sperm which have died because they ran out of life, and sperm which left the simulation and swam into the end of the ampulla.

The [Movement State Timer] and [Remaining Life], as defined in Table 4.2, are

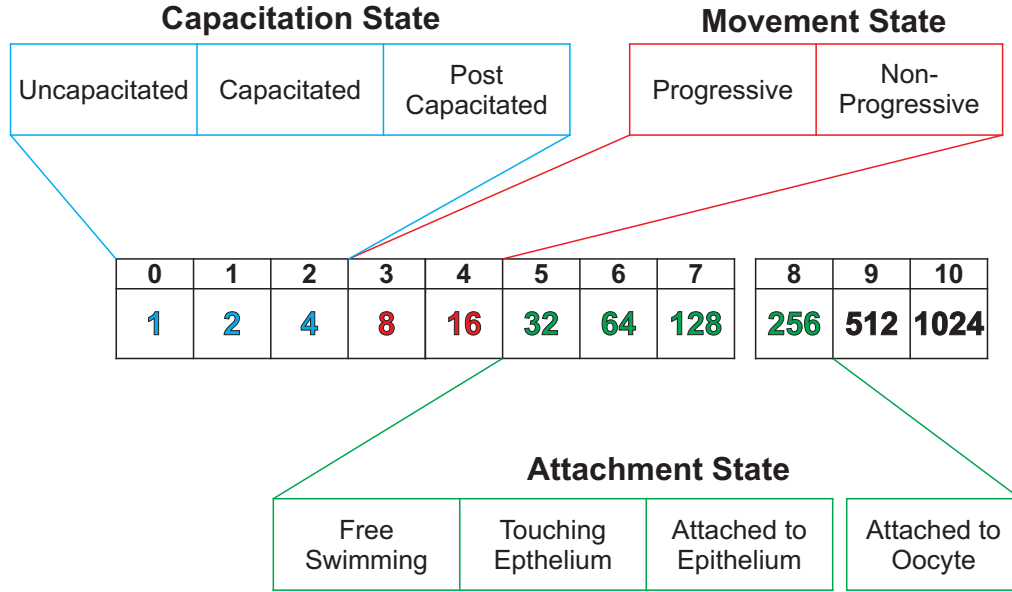


Figure 6.3: How the capacitation state, movement state, attachment state and alive state are encoded into a single variable. The alive state is implicitly determined by the presence or absence of the capacitation state.

integers ([Movement State Timer] and [Remaining Lifetime] in Table 6.2).

To simplify tracking and analysis of the model, several additional properties were defined that are not listed in Table 4.2. Each sperm agent contains a unique identifier ([ID]), which allows the state of each agent to be tracked over time. The agents also contain an [Oviduct Segment] parameter, which holds the current position of the agent along the oviduct tube relative to the individual segments defined in Section 5.7.1. The agents contain two additional properties which are used once the sperm has attached to an oocyte. These properties are the ID of the oocyte it is attached to ([Attached Oocyte ID]) and the simulation time (s) in which the attachment was made ([Attached Oocyte Time]).

Oocyte agents are relatively simple. The [Current Position], as defined in Table 4.2, is encoded as a 3 element float array ([Current Position] in Table 6.2). Oocyte agents also have a unique identifier ([ID]) to allow tracking of which oocyte each sperm attaches to.

### 6.2.3 Agent functions

The functionality of the conceptual model is encoded as a series of functions. Despite the oocyte agent only playing a passive role in the simulation, it still has functions and raises messages. As the position and direction of the sperm agents are encoded as transformation matrices, a set of functions for matrix manipulation was written. The probability mechanisms used in the simulation are implemented as a simple probability test. A random number is calculated between 0 and 1, and if the random number is lower than the probability threshold then the outcome is true.

The two other main functions used within the simulation are the conic rotation function described in Section 4.3.2 and collision detection between sperm agents and the environment.

#### 6.2.3.1 Conic rotation implementation

The conic rotation is applied when sperm are reflecting or detaching from the oviduct wall, or when they are turning due to non-progressive movement. Initially, a direction of rotation is established. A point around the perimeter of a circle is randomly chosen, which determines the direction of rotation relative to the `[Current Direction]`. Then the magnitude of rotation is randomly chosen between zero and the maximum angle of rotation (`[Max. Non-Progressive Angle]`) system level parameter. The relative pitch and yaw rotation angles are then calculated and applied to the `[Current Direction]`.

#### 6.2.3.2 Collision detection implementation

Interactions between sperm and the oviductal environment can have a significant impact on sperm navigation and behaviour. In order to model these interactions computationally, an algorithm was devised to efficiently detect collisions. Many collision detection strategies combine multiple techniques, which are selected depending on the shape, size, complexity and behaviour of objects within a 3D environment. Collision detection strategies can be split into two categories, discrete and continuous. Discrete strategies check for collisions at regular, discrete time points. The accuracy of this approach relies on a large number of small time steps. If the distance between time steps is too large, then two objects can potentially pass through each other if moving with sufficient velocity. Continuous collision detection is a way of tracking collisions at any moment in time. This strategy removes the possibility of two objects passing through each other. However, it is more complicated and computationally intensive.

A naïve collision detection implementation will perform intersection tests between a

moving object and every triangle in the environment, resulting in  $O(n^2)$  triangle intersection tests, where  $n$  is the number of triangles. The majority of collision detection algorithms aim to reduce the number of intersection tests as much as possible using a combination of broad phase and narrow phase techniques. Broad phase techniques involve identifying large groups of triangles which cannot possibly intersect. Narrow phase techniques then process the remaining triangles one at a time to determine if a collision has occurred.

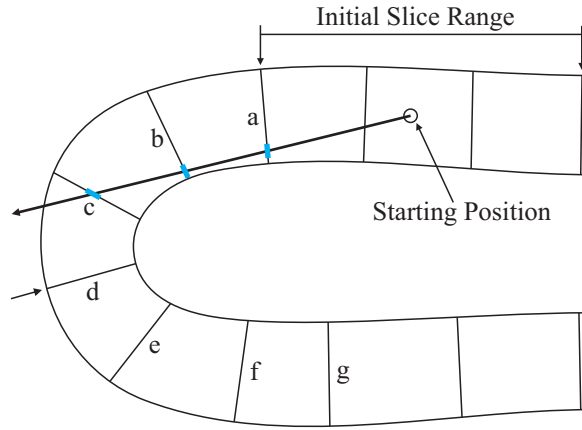
One method commonly used to reduce the number of intersection tests is scene partitioning. The space in which objects exist within the 3D environment is split into sections and only objects within each section are tested for intersections. Various partition methods exist, such as Octrees (Moore and Wilhelms, 1988), Binary Scene Partition (BSP) trees (Thibault and Naylor, 1987) and Voxel Grids (Garcia-Alonso et al., 1994), which group the objects together in different ways based on their spatial position.

Another efficient method is through the use of bounding volumes. A complex object is wrapped in a simple object and collision tests are performed between the bounding objects to eliminate pairs of objects that are too far apart to collide. Where the bounding volumes of two objects do collide, a per-triangle intersection test is then performed between the underlying objects to determine if any collision has taken place (Watt and Policarpo, 2000).

As sperm move very fast in relation to their size every iteration, a continuous collision detection strategy was chosen. The chosen strategy was a Sphere Sweeping algorithm (Rouwé, 2003), modified to run in CUDA. A sphere representing each agent is projected along its intended path and the position along the path where the sphere intersects with each triangle is detected. For each triangle, the current sphere position and direction are used to determine if the sphere can collide with the plane that the triangle lies on. If an intersection is possible, then the position and direction are converted into 2D space, relative to the triangle plane. The circle formed by the projection of the sphere in 2D is then tested for intersection with the triangle (Rouwé, 2003). As the triangle intersection test is computationally expensive, the number of intersection tests performed each iteration should be reduced as much as possible. Therefore, a combination of scene partitioning and bounding volumes were used to minimise the number of intersection tests.

As the 3D environment is essentially a long tube, it can be split into sequential slices. Each sperm agent tracks its current position along the tube. During movement, each agent determines the slices with which it could potentially collide. The line formed

between the current position and the potential new position, which is calculated using the direction vector and the movement distance, is used to perform a series of intersection tests with the plane marking the edge of each slice. From the starting position, slices are sequentially tested in-front of the current slice and then behind the current slice. The process stops if the plane from a slice is not intersected. This is illustrated in Figure 6.4. The initial slice range is set to one slice before and one slice after the current slice. Intersections with the slice planes are progressively checked from (a) to (g). The movement path does not intersect with slice (d) so the algorithm interrupts the process and marks the outer boundary as slice (d).



*Figure 6.4:* How the potential slices are identified. The initial slice range (3 slices) is progressively extended as intersections with the corresponding slice planes are identified. The algorithm continues until slice plane (d) is processed where no intersection occurs. The final slice range is 6.

Reducing the number of triangles by slice identification significantly reduced the number of triangle intersection tests, but this was further reduced by wrapping each triangle in a bounding sphere. Before performing the triangle intersection test, the distance between the start point and the triangle bounding sphere was tested. If it was larger than the maximum moving distance, then the triangle intersection test was not needed. If it was within the maximum moving distance, and could therefore potentially collide, then a line-sphere intersection test was performed. If the line intersected the sphere, then finally the triangle intersection test was performed.

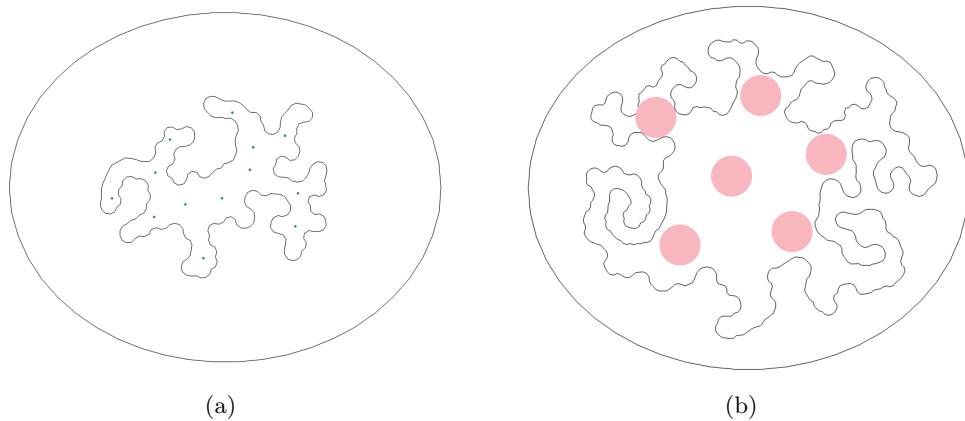
3D models are typically collections of vertices within 3D space, with triangle definitions linking vertices, resulting in a set of surfaces. As the collision detection algorithm performs intersections in 2D space relative to each triangle, it was inefficient to store the model in a traditional manner. For optimisation with the collision detection algorithm,



instead of storing the triangle and vertex positions in 3D space, the x and y position of the vertices forming each triangle in 2D space was pre-calculated and stored in the primary data structure. The collision detection algorithm uses the triangle plane (normalised vector + w offset), the vector perpendicular to the plane (U) and the cross product of the plane normal and U to convert between 3D space and 2D space local to the triangle. These values were also pre-calculated and stored in the primary data structure. This reduced the number of calculations required each iteration. A bounding sphere for each triangle was pre-calculated and used for the initial intersection tests. At the start of the simulation, the data structures containing the full definitions for the 3D environment were loaded onto the GPU as texture memory, which is optimised for read only access.

### 6.3 Model initialisation

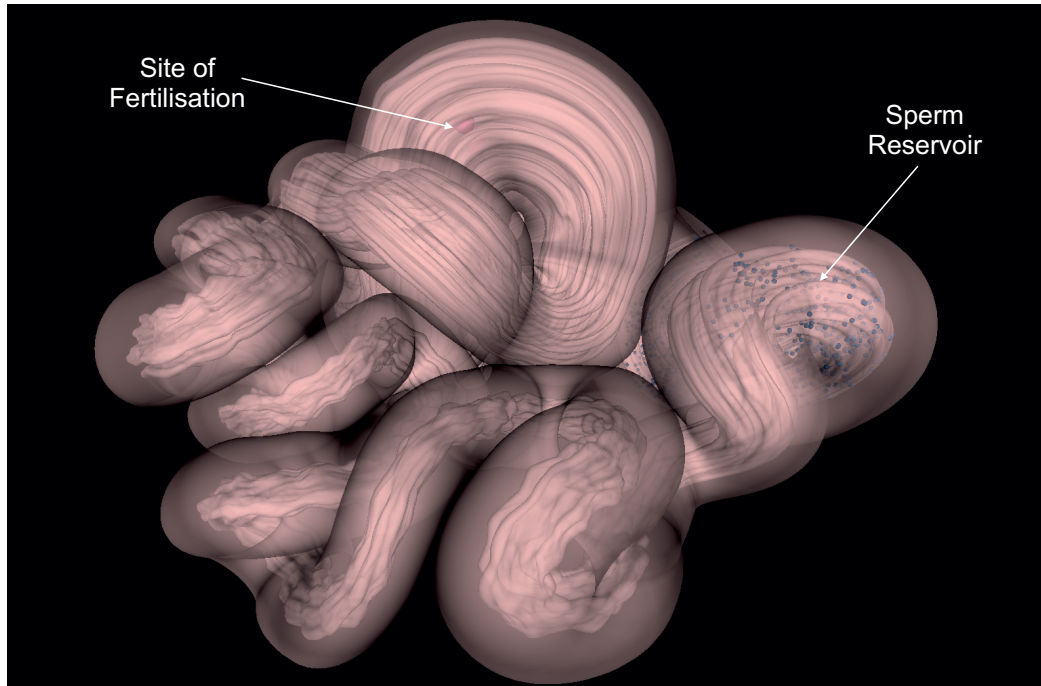
This section describes how the initial agent position is determined for all simulation runs in Chapters 7 and 8. The initial position of sperm agents is randomly scattered along the oviduct wall of the most intramural and extramural segments of the oviduct. Using the slices which were identified as being in the caudal isthmus, a set of initial positions within the lumen were identified, as shown in Figure 6.5(a). For oocyte agents, the set of initial positions are identified within the slices which were identified as being at the site of fertilisation, as shown in Figure 6.5(b). The initial direction of sperm agents was randomly determined within a  $60^\circ$  cone relative to the direction of the tube. This ensures that all sperm were initially pointing in the correct direction, as would be the case if they had just entered the tube from the uterus.



*Figure 6.5:* (a) An oviduct cross-section with possible initial positions for sperm agents and (b) an oviduct cross-section at the site of fertilisation with possible initial positions for oocyte agents.

When a new set of agent initialisation files is generated, the position of the sperm and oocyte agents are randomly chosen based on the identified initial positions. For oocyte agents, the initial position is the position they will stay in for the whole simulation.

Sperm agents should be initially attached to the wall of the oviduct. As the initial positions were within the lumen, a pre-initialisation step runs at the beginning of the simulation to distribute the sperm agents against the nearest oviduct wall. The algorithm that performs the initial distribution against the oviduct wall is deterministic, and the direction in which sperm agents are moved so that they are in contact with the wall is calculated based on the agent id. This ensures that the starting position will be the same on subsequent simulation runs using the same initialisation file. The initial position of the sperm agents and oocytes in the model are shown in Figure 6.6.



*Figure 6.6:* The 3D oviduct model (OvA1) that was used for the initial simulations. The initial position of sperm agents within the sperm reservoir and the oocytes at the site of fertilisation are shown. The displayed size of the sperm agents has been increased for visual clarity.

## 6.4 Outputs

This section describes the outputs that are generated by the model and how the data is managed, analysed and visualised.

### 6.4.1 Data management and analysis

Typically, agent based simulations record the position and state information for each agent every iteration. When run for a simulated 8 hour period (28,800 iterations) with 1,536 agents and multiple replicates, this could potentially result in a massive amount of data.

The FLAME GPU framework generates data in XML format, and the initial 0.XML file for a 40 replicate simulation was 47MB, resulting in a potential total results set of 1.3TB per simulation run, which is clearly unmanageable. The framework was modified to store the simulation data in binary format, with an initial header file describing the structure of the data, and then the iteration records following sequentially. The location of data for each iteration was stored in a lookup table in the file header. This significantly reduced the potential final simulation data size down to 130GB. The framework was further modified to only export data at specific intervals, in this case every 1,800 iterations, representing every 30 mins. As the observational data in biological studies is usually oriented around hourly observational measurements, this provided sufficient data for the analysis. It also had the advantage that all the data required for statistically significant analysis of a single simulation run to be performed was stored in a single output file, with a size of around 40 MB. The data generated by the framework was aggregated to store the number of sperm per oviduct slice with respect to time. Each slice was mapped onto a corresponding region of the oviduct, such as the isthmus or the ampulla, so the number of sperm in each oviductal region at specific time points could easily be established. The aggregated data in each file was imported into a Microsoft SQL Server database for analysis.

Several different queries were written to generate different views of the distribution of sperm within the oviduct in relation to the oocyte.

- The average number of sperm attached to oocytes at regular intervals.
- The average number of oocytes with sperm attached at regular intervals.
- The average time for first sperm to reach the an oocyte.
- The average number of sperm with specific states in each region at regular intervals.

### 6.4.2 Visualisation

In order to present the simulation to biology experts, it was necessary to construct a tool to allow visualisation of the data. This tool had three main requirements:

- 1 Ability to see a smooth animation of the simulation despite the large file size
- 2 Visualisation of the 3D environment
- 3 Visualisation of agents movement over time

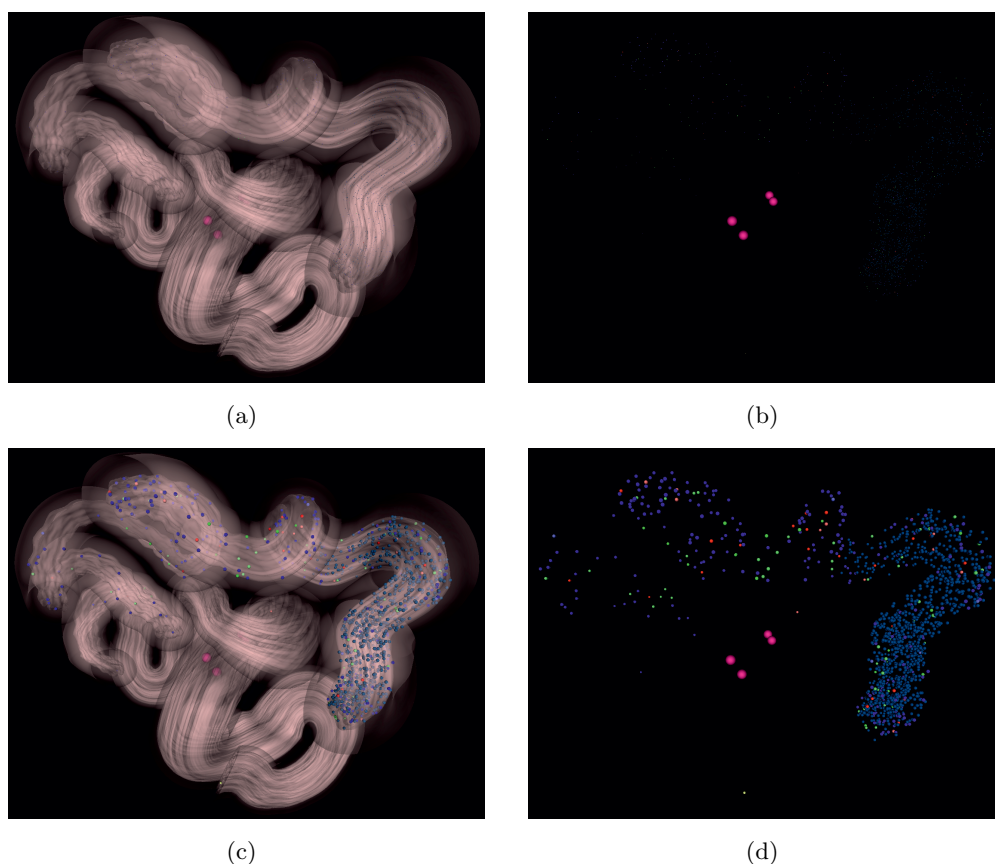
For the first requirement, the aggregated data structure described in Section 6.4.1 is not suitable for this type of display. Instead, to visualise the interactions, the simulation data was recorded using a single replicate, recording the position of all agents every iteration. Each iteration the visualiser is able to read the chunk data of the next iteration and link the agents by their identifiers. Therefore, only the currently displayed records are retained in memory, and all other data remains in the binary data file.

For the second requirement, the 3D environment is highly complex, with over 600,000 triangles. The 3D models used for the simulations were saved in a triangle strip based data format, which is optimised for 3D display. In order to be able to view the interactions that occur within the model, the visualiser required the ability to see through the 3D environment. Therefore, the ‘dual depth peeling’ technique developed by Bavoil and Myers (2008) was modified to run under CG and implemented in the visualiser. This technique allows the inside and outside of the tube to be visualised along with the interactions which are occurring within.

For the third requirement, the agents were displayed as simple spheres with a relative radius matching their size in the agent model. The colour of each agent is dependent on its current state. Figure 6.7(a) shows the 3D model with transparency applied and Figure 6.7(b) shows the same scene with the 3D model completely hidden. The position of the oocyte agents (large pink spheres) are visible in both images, but the sperm agents are more difficult to see.

As sperm agents are very small in relation to their environment, an option was added to the visualiser to increase the radius of the displayed spheres to make them easier to see. Figure 6.7(c) shows the 3D model with transparency applied with enhanced radius sperm agents and Figure 6.7(d) shows the same scene with the 3D model completely hidden. The position and state of the sperm agents is now easier to see in both images.

As sperm agents move a long distance in relation to their size every iteration, on playing the simulation the agents appeared to jump around the environment and it was not clear how they were moving. To address this problem, a line was drawn for each agent connecting the previous four positions to the current position. This line allowed the path and progress of each agent to be clearly tracked through the environment. It also had the unintended side effect of making the agents look more



*Figure 6.7:* (a) The environment with transparency applied and the agents visible through the structure, (b) the agents displayed without the environment, (c) the environment with transparency applied and the radius enhancement of sperm agents and (d) the radius enhancement of sperm agents without the environment. oocyte agents.

like sperm, as the path looked like a tail. Figures 6.8(a) and 6.8(b) show the impact of the addition of a connecting line on the simulation display with and without the environment with normal sized agents, and Figures 6.8(c) and 6.8(d) show the impact of the addition of a connecting line on the simulation display with and without the environment with the enhanced radius agents. These images clearly show the benefit of the display modifications. When played as an animation, the direction and nature of the movement of agents was clearly seen. Having a good visualiser was key to the success of this project, as it made it very easy to demonstrate to biologists and non-computer scientists exactly how the simulation was running and what a simulation run looks like. This facilitated group discussions and provided a means of gaining meaningful feedback from experts in the biology domain, which was invaluable for model development, as discussed in Chapter 4.

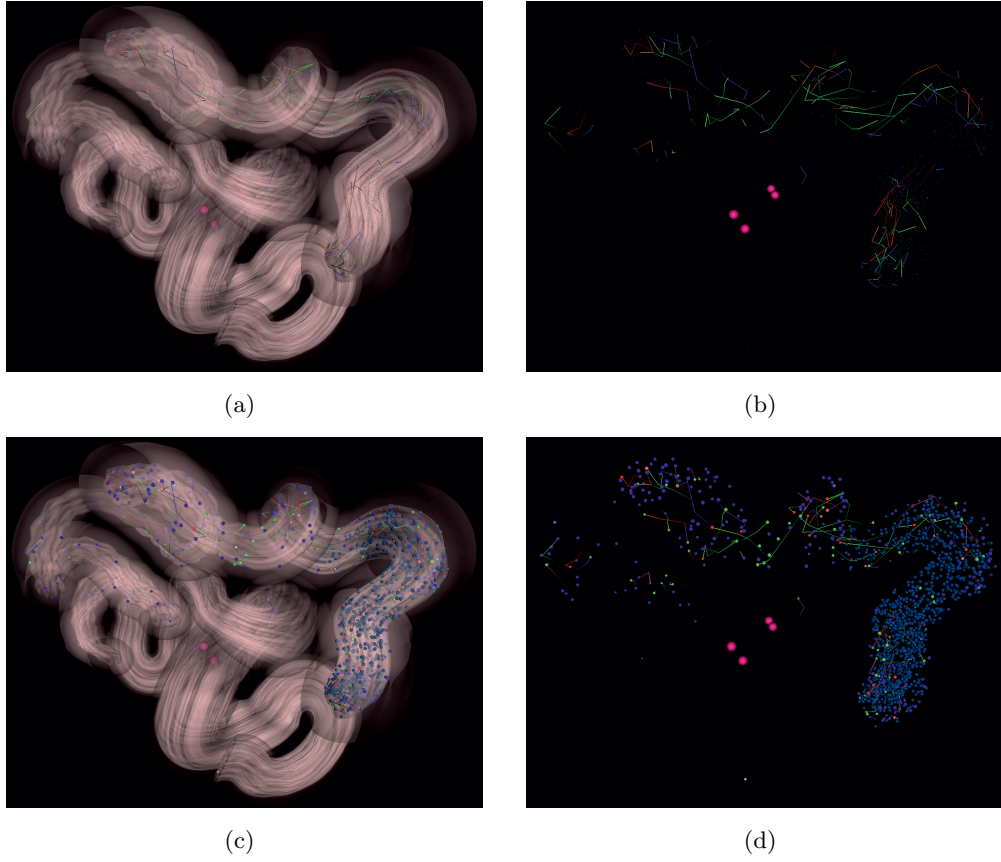


Figure 6.8: (a) The environment with transparency applied and the agents with lines showing their path, (b) the agents and lines showing their path displayed without the environment, (c) the environment with transparency applied, the radius enhancement of sperm agents and lines showing their path and (d) the radius enhancement of sperm agents and lines showing their path without the environment. oocyte agents.

## 6.5 Optimisation

This section describes the performance results due to optimisation applied to the collision detection strategy and through the use of concurrent simulations. The hardware configuration used for all simulations is shown in Table 6.3.

Hardware Configuration
Windows 7 x64
Intel Core 2 Quad Q6700 @2.66Ghz
4GB RAM
NVIDIA GeForce GTX 280 [1GB]

Table 6.3: Hardware configuration used for all simulations.

### 6.5.1 Collision detection

A set of simulations were run in order to test the significance of the collision detection optimisation described in Section 6.2.3.2. Each simulation contained 1,536 agents, using one of the oviduct models which is made up of 648,930 triangles split into 1,200 slices. Each simulation ran for 300 iterations, representing a simulation time of five minutes. Figure 6.9(a) shows the speed of the naïve collision detection algorithm, along with the inclusion of scene partitioning, bounding spheres and the fully optimised solution which contained both. The naïve algorithm took over 83 minutes to complete 300 iterations, whereas the optimised algorithm took just over 9 seconds to complete the same number of iterations. Due to the massive difference in time, a logarithmic scale is shown for comparison. As a full simulation could potentially run for over 28,800 iterations, the value of collision detection optimisation becomes clear.

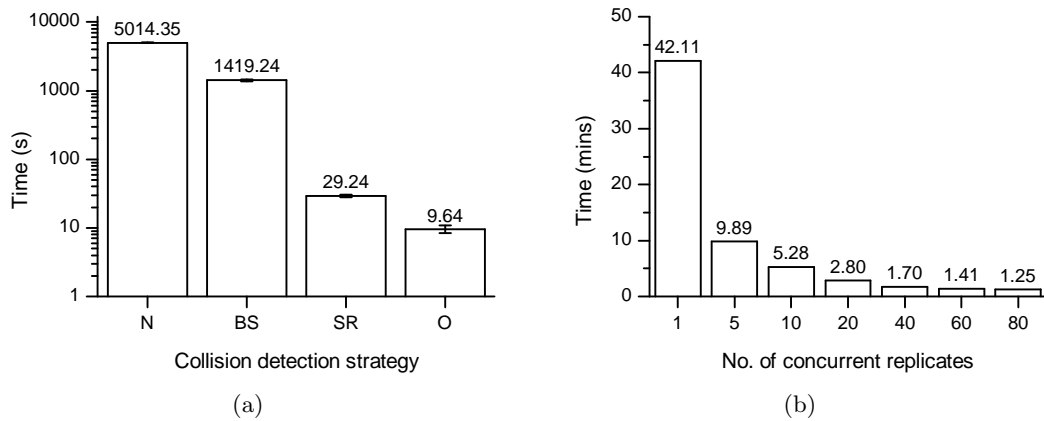


Figure 6.9: (a) The time taken to complete 300 iterations using different collision detection strategies. The scale is logarithmic due to the exponential difference in time between the naïve and optimised approaches. N = naïve, BS = bounding sphere, SR = slice range, O = optimised. (b) the average time taken to complete a single simulation when multiple simulations are run concurrently.

### 6.5.2 Concurrent simulations

FLAME GPU is highly optimised for processing very large numbers of agents concurrently (Richmond and Romano, 2011). As this simulation only requires a relatively small number of agents, around 1,500, the model was modified to allow multiple simulations to run concurrently. All agents run within the same 3D model, but are conceptually in different simulations, identified by a unique environment id. Only sperm agents and oocyte agents with matching environment ids can interact with each other. This means that multiple replicates of the simulation can be run concurrently. To

demonstrate the benefit, the time taken to run a simulation with 1,536 agents for 28,800 iterations, with different numbers of concurrent replicates was recorded.

Figure 6.9(b) shows the average time for each simulation to complete when different numbers of replicates are run concurrently. The average time is calculated as the total run time divided by the number of concurrent simulations. The results show that when 10 simulations are run concurrently, the average time per simulation is significantly lower than a single simulation. When the number of concurrent becomes greater than 40, then the maximum capacity of the GPU is achieved, and the relative performance increase slows down, with each simulation taking approximately one minute to complete.

## 6.6 Conclusions

The main aim of this chapter was to describe the construction of an agent based model of sperm movement within the oviduct. An evaluation of agent based modelling frameworks was carried out, and FLAME GPU was identified as the most appropriate modelling framework for the application. The methodology of the computational realisation of the conceptual model has been described, along with the optimisation of the collision detection algorithm. Managing the massive amount of data generated by the agent simulation, and how the data was analysed and visualised was presented. The parallel processing power of the GPU has been harnessed to run multiple concurrent simulation replicates, meaning that a single model run generates sufficient data to infer statistically significant results. The GPU has also been used to perform computationally intensive visualisation of the simulation results within the 3D environment. Chapter 7 describes calibration, parameter analysis and validation of the model.



## Chapter 7

# Calibration and validation

This chapter describes validation of the computational model of sperm behaviour. First, literature surrounding validation of computational models is discussed, then the process used to validate this model is described and literature used to calibrate the model is summarised. A set of initial simulations were run to identify the significance of variations in the estimated and uncertain parameters used in the model. The model is carefully calibrated to ensure the number of sperm reaching the site of fertilisation over time is not significantly different from the values reported in literature. Once calibrated, sensitivity analysis is performed on the parameters of the model. Finally, the simulation is validated against two independent studies which were not used to calibrate the system, and the required steps to fully validated the model are discussed.

### 7.1 Validation

For a model to be useful, it is necessary to prove that the way the model behaves is an accurate representation of the real world system. This will allow aspects of the model that have not been investigated through biological experimentation to be tested and will give confidence in the results generated by the model. Validation techniques can be employed to ensure that the model is correct, and that assumptions and simplifications used to create the model do not have a significant impact on the generated results.

The validity of a model is the degree to which the conceptual model accurately recreates the behaviour of the real world system (Kleijnen, 1995). A model is not an exact reconstruction of the system being investigated, but is an abstraction of that system, usually focusing on a specific aspect (Kleijnen, 1995). However, this does not mean the model is not capable of making accurate predictions about the real system (Carley, 1996).

A large number techniques, frameworks and approaches have been described to guide

model development and validation (Klügl, 2008). These are typically multi-step processes, running concurrently with model development. Different aspects of the model are validated at every step, and the models are recursively refined until an adequate validation criteria is reached (Kleijnen, 1995; Klügl, 2008; Law, 2005; Sargent, 2007). Where the parameters of a model are estimated or uncertain, then model calibration or tuning is often performed. This is where the values of uncertain parameters are systematically varied until the results match a target set of data (Carley, 1996; Thorne et al., 2007). Validation of the model can then be performed using different, independent sets of data.

The ideal approach to modelling biological systems is to create the model based on experimental data. For a given real world problem, a conceptual model is developed. A computational realisation of the conceptual model is then created and tested to ensure that it is an accurate representation of the conceptual model. The computational model can then be experimented with to validate it against the real-world problem, and new real-world experiments can be performed. The results of this experimentation can be used to update the conceptual model, and then the computational model. This process is repeated until an accurate representation of the real world problem is established (Ideker et al., 2001; Sargent, 2005). This approach is not always feasible, due to lack of resources and availability of expertise from different fields.

### 7.1.1 Levels of validation

Computational models can have different levels of validation. At the lowest level, face validation, which encapsulates aspects of individual and emergent behaviour, can be obtained if the reasonableness of assumptions can be established and the resulting simulation appears correct. This is often a very subjective process, based on discussions with experts on the system being modelled (Klügl, 2008). One approach proposed by Klügl (2008) is to show animations of the results showing the group behaviour and individual behaviour to experts in the field and obtain verbal agreement that the observed individual behaviour and the pattern of population level behaviour is plausible.

This can be further enhanced by several factors. If the model is grounded in reality (i.e. the spatial and temporal scales used within the model are the same as the real system) then comparisons can be made with the real system along the same scales. If the processes and parameters used within the model are the same as those in the real system, then the individual processes of the model can be compared with the known behaviour of the individual processes of the real system. If the values used to parameterise the model hold the same value or are within a range of values based

on accurate measurements of the real system, then a direct comparison between the inputs and outputs generated by both the real system and the model can be made. Finally, statistical validation of the model output can be performed to ensure that the results generated by the model are a realistic representation of the real world system when given the same inputs. When comparing model generated outputs and biological outputs, care has to be taken to ensure that the two types of data are comparable. For example, the same types of measure should be made from both systems with the same degree of accuracy. The type of statistical analysis performed depends on the type of data available. Section 7.1.5 describes the type of statistics used in this thesis.

### 7.1.2 Sensitivity analysis

Mathematical techniques such as sensitivity analysis are available to help investigate the behaviour of a model in the absence of real data. Sensitivity analysis is a technique that can help to establish the sensitivity of the model to variations in individual parameters (Saltelli et al., 2008). A mathematical relationship is established between variations in input parameters and the output of a model. The parameters of a model are systematically modified by a fixed ratio, and the impact on the results of the model are then determined. For example, if the value of a single parameter is increased by ten percent, and the results are altered by fifty percent, then the parameter can be considered to have high sensitivity to change. Alternatively, if the results are only altered by two percent, then the parameter can be considered to have low sensitivity to change. Highly sensitive parameters could indicate areas of the system that require more detailed or accurate modelling, or areas that would benefit from experimentation with the real world system. Parameters with low sensitivity could possibly be abstracted away, or set to a fixed value, thus simplifying the model (Kleijnen, 1999).

### 7.1.3 Validation process

Validation of the agent based model of sperm behaviour described in Chapter 4 was performed using a multi-step process, which is summarised in Figure 7.1. First, simulations were run using the initially identified parameters. Then, the maximum potential impact of the parameters which were estimated or marked as uncertain was established. The model was then calibrated by systematically modifying the values of the uncertain parameters until the results were statistically similar to those reported in literature. Once calibrated, the sensitivity of the model to variations in individual parameters was established, along with the impact that modification to the 3D environment to compensate for histology shrinkage had on the results. Finally, the model was statistically validated against two different biological studies.

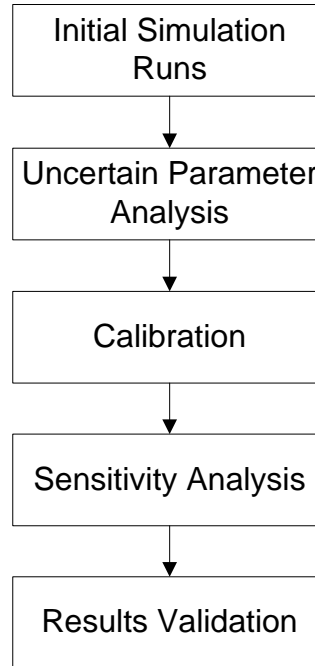


Figure 7.1: The processes followed for calibrating and validating the model of sperm behaviour.

#### 7.1.4 Literature for calibration

As experimental data was not available for this project, a set of data from literature was identified for model calibration. The decision was taken to calibrate the model based on the distribution of sperm, more specifically the number of sperm reaching the site of fertilisation in the ampulla. Three studies were identified which provide counts of the number of sperm in the ampulla of the mouse after 9 hours *post coitus*, which were Bennett and Dunn (1967), Braden (1958) and Braden and Austin (1954*b*). In all these studies, the same method was used to determine the number of sperm at the site of fertilisation. The wall of the distended region of the ampulla was torn or cut open, and the contents allowed to flow into a dry petri dish. The oviduct tissue was then removed from the petri dish and the contents inspected. The number of sperm in the ampulla region are used for comparison. Sperm in the pre-ampulla and ampulla-isthmus junction are assumed to be too far away from the site of fertilisation to be included. See section 5.7.1 for a breakdown of the individual regions of the model.

The results from these studies are summarised in Table 7.1. The differences between the counts for these studies are due to different strains of mice and different cross-strain

combinations (Braden, 1958). The pooled mean value of 21 sperm in the ampulla was calculated from the studies in Table 7.1 and used as a guideline for calibration.

Mean no. of Sperm	Range	No. of Oviducts	Source
$27.9 \pm 46.6SD$	0 - 208	32	Bennett and Dunn (1967)
$23.7 \pm 16.6SD$		41	Braden (1958) from CBA/CBA
$27.2 \pm 19.2SD$		34	Braden (1958) from CBA/C57
$16.6 \pm 20.8SD$		18	Braden (1958) from RIII/C57
$14.8 \pm 15.9SD$		22	Braden (1958) from RIII/A
16.9	2 - 75	30	Braden and Austin (1954b)

Table 7.1: The expected total number of sperm at the site of fertilisation after 9 hours *post coitus*. The multiple measurements from Braden (1958) are from crosses of different strains of mouse.

### 7.1.5 Metrics for calibration and validation

Unless otherwise stated, simulations were run using a single 3D environment with 40 concurrent replicates over a simulated 8 hour period, with measurements recorded at 30 minute intervals. As the simulation starts 1 hour *post coitus*, the end of the simulation will represent 9 hours after insemination. Natural mating is assumed in all simulations. This means that insemination occurs before ovulation, as described in Section 4.3.2.5. Several different measures of the system were used during model calibration:

- The number of capacitated sperm at any moment in time within the entire oviduct.
- The total number of sperm reaching the site of fertilisation at the end of the simulation.
- The final distribution of sperm at the end of the simulation.
- The number of capacitated sperm at any moment in time in the ampulla.
- The average number of sperm reaching the site of fertilisation over the simulation.

Initial analysis of the data indicated that the variations in the distribution of sperm and oocyte penetration rates between individuals did not follow a normal distribution. Therefore, when two sets of data being compared have individual data points, then statistical significance was determined using a non-parametric wilcoxon signed-rank test. However, some of the data in literature is provided as pooled totals. In these

cases, significance is determined using the Chi Square test. Due to the large number of replicates within the simulation, significance for all tests was determined at  $p=0.001$ .

## 7.2 Model calibration

This section describes the model calibration process.

### 7.2.1 Initial simulation runs

The first step of the validation process was to run the simulation using the initially identified parameters, as described in Chapter 4. Figure 7.2(a) shows the final distribution of sperm in different oviductal regions at the end of the simulation and Figure 7.2(b) shows the percentage of the total population which were capacitated at different time points. The results show that no sperm were able to reach the ampulla or pre-ampulla regions of the oviduct within the simulation period. As we know that the mean number of sperm within the oviduct after 8 hours must be greater than 0, this indicates that the initial values for the simulation parameters are incorrect. As several of the parameters are marked as uncertain and do not have accurate values from literature, the potential impact of variations in these parameters were investigated.

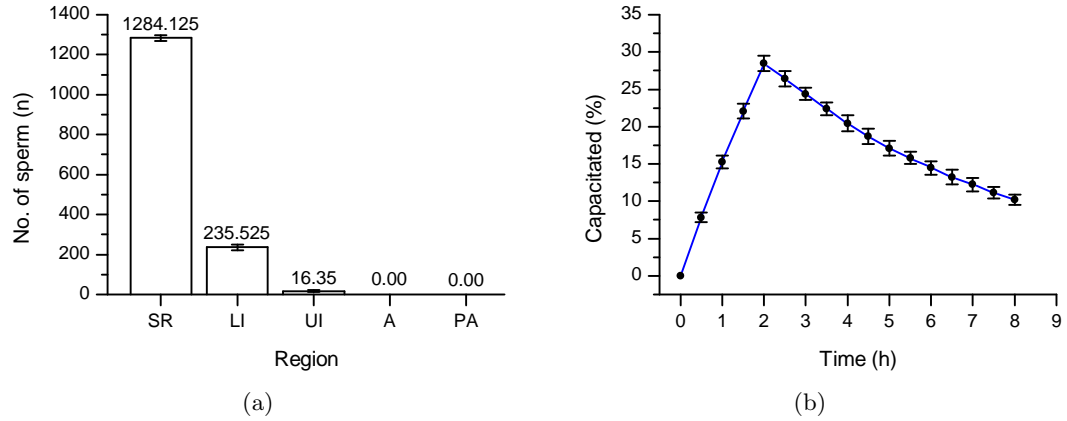


Figure 7.2: (a) The final distribution of sperm when run with the initial simulation parameter values and (b) the percentage of sperm which are capacitated over time. Error bars show S.D. SR = sperm reservoir, LI = lower isthmus, UI = upper isthmus, A = ampulla, PA = pre-ampulla.

### 7.2.2 Uncertain parameter analysis

This section describes the process followed to identify more realistic values for the parameters where the data from literature is either not available or not reliable. The

chosen approach was to investigate the potential impact of each parameter individually by varying it through its full range. A value can then be chosen which gives a more realistic output. Although more complex analysis methods such as Latin Hypercube analysis (Reed et al., 1984) are available for parameter estimation, the simple approach was fast and resulted in a reasonable set of parameters.

The simulation parameters which were estimated or marked as uncertain in Chapter 4 are the [Capacitation Threshold], [Progressive Attachment Threshold], [Max. Reflection Angle], [Max. Detachment Angle], [Max. Non-Progressive-Angle] and the [No. of Movement Steps]. These parameters are involved in different processes. The [Capacitation Threshold] is involved in the [TRY TO CAPACITATE SPERM] process, which governs the rate of release from the sperm reservoir, and therefore the total number of active sperm at any moment in time.

The other uncertain parameters, namely the [Progressive Attachment Threshold], [Max. Reflection Angle], [Max. Detachment Angle], [Max. Non-Progressive-Angle] and [No. of Movement Steps], are all related to interactions with the oviduct, and control the distribution and progression of sperm within the environment over time. As these processes are logically different, they can be investigated in isolation.

Variations in the [Capacitation Threshold] will influence the distribution and progression of sperm within the environment, but variations in the other uncertain parameters will not influence the number of capacitated sperm. Therefore, the [Capacitation Threshold] was investigated first.

### 7.2.2.1 Capacitation

The purpose of this simulation is to identify the range of values which result in a change in the simulation behaviour. The [Capacitation Threshold] controls the rate at which sperm are capacitated throughout the simulation. It is a probability, and therefore has a maximum potential range of between 0 and 1. A set of simulations was run to investigate the [Capacitation Threshold] through its entire range, varying it between 0 and 1 in 0.1 intervals. Figure 7.3 shows the impact that variations in the capacitation threshold have on the percentage of the sperm population which are capacitated at any one time.

This graph shows that when the [Capacitation Threshold] is zero, no sperm are capacitated. For all values of the [Capacitation Threshold] greater than zero, all sperm become capacitated in the first 30 minutes, and all are dead by 2.5 hours. These results indicate that in order for there to be a gradual release of sperm over time, the

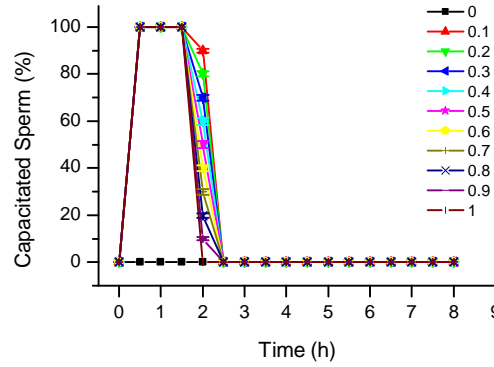


Figure 7.3: Analysis of the [Capacitation Threshold] when varied between 0 and 1 in increments of 0.1.

threshold must have a value between 0.0 and 0.1.

Considering that the initial value for the [Capacitation Threshold] was based on the time for all sperm to become capacitated, the parameter was systematically varied using the formula specified in Section 4.3.2.1, reproduced below for convenience.

$$P = \frac{1}{(ct \times 3600)} \quad (7.1)$$

The value of (ct) was varied between 1 and 24, in increments of 1. The results, which are expressed as the percentage of capacitated sperm over time, are shown in the heat map in Figure 7.4(a). At T=2 hours, the percentage of capacitated sperm for all investigated values is at the peak. This corresponds to the lifetime of capacitated sperm, which is also 2 hours. After this time, sperm start to die, and the result is a gradual decline in the percentage of capacitated sperm.

The percentage of capacitated sperm in mice is not known. However, in humans, it has been observed to vary between 2% and 14% (Cohen-Dayag et al., 1995), with an overall average of 10% (Jaiswal and Eisenbach, 2002). This data was used to select an appropriate value for the [Capacitation Threshold].

Figure 7.4(b) shows the percentage of capacitated sperm at T=2 hours in relation to the [Capacitation Threshold]. The threshold value 14%, which is shown on the figure as a red line, is used as the maximum. This gives a capacitation threshold of ct=13. When looking at the data in a tabular form, the average over the entire time period for ct=13 is 9.9%, which matches the average overall capacitation rate of (10%) provided in Jaiswal and Eisenbach (2002). This indicates that choosing a value of ct=13 is a



reasonable assumption.

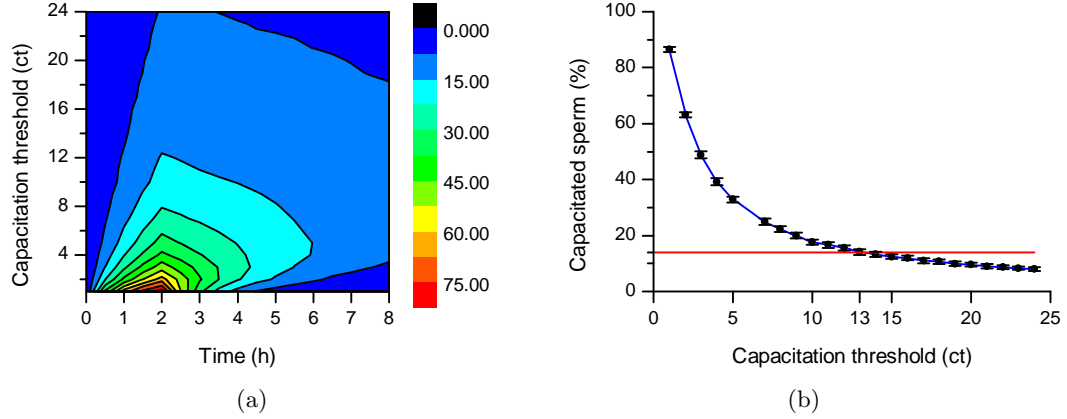


Figure 7.4: (a) A heat map of the [Capacitation Threshold], varying (ct) between 1 and 24 in increments of 1 and (b) how the percentage of capacitated sperm at  $t=2$  hours changes as the [Capacitation Threshold] (ct) is varied between 1 and 24 in increments of 1.

#### 7.2.2.2 Progression and distribution

Once a reasonable value for the [Capacitation Threshold] was established, then the variables involved in sperm distribution and progression were investigated. An initial set of simulations was run for the remaining uncertain parameters to identify which parameter has the largest potential impact on the results. A single simulation with 40 replicates was run for each parameter combination.

The [Progressive Attachment Threshold] controls the likelihood of a sperm agent attaching to the oviduct when a collision is made. It is a probability, and therefore has a maximum potential range of between 0 and 1. The [Progressive Attachment Threshold] was varied through its full range between 0 and 1 in 0.1 intervals.

The [Max. Reflection Angle] controls the maximum angle that sperm can reflect off the oviduct, the [Max. Detachment Angle] controls the maximum angle that sperm can detach from the oviduct once attached and the [Max. Non-Progressive Angle] control the maximum angle that sperm can rotate when moving non-progressively.

All three parameters are angles, with a full range of between 0 and 180 degrees. They were varied through their full range between 0 and 180 degrees in 20 degree intervals.

The [No. of Movement Steps] controls the maximum potential collisions that a sperm agent can perform during a single iteration. It is a whole number, and has a potential range of between 1 and 12. The [No. of Movement Steps] was varied

between 1 and 12 in intervals of 1.

Figure 7.5(a) shows the full potential range when the data for all values of each parameter are pooled. This shows that a small number of values for the [Progressive Attachment Threshold] can result in the largest number of sperm in the ampulla, but the majority of values are close to zero. The [Max. Non-Progressive Angle] can have a smaller impact on the number of sperm in the ampulla. The remaining parameters have very little impact on the number of sperm in the ampulla.

Figure 7.5(b) shows the potential impact of each parameter on the number of sperm in the ampulla after 8 hours, assuming all other parameters have their initial values. The potential impact was identified as the largest number of sperm in the ampulla for all parameter values.

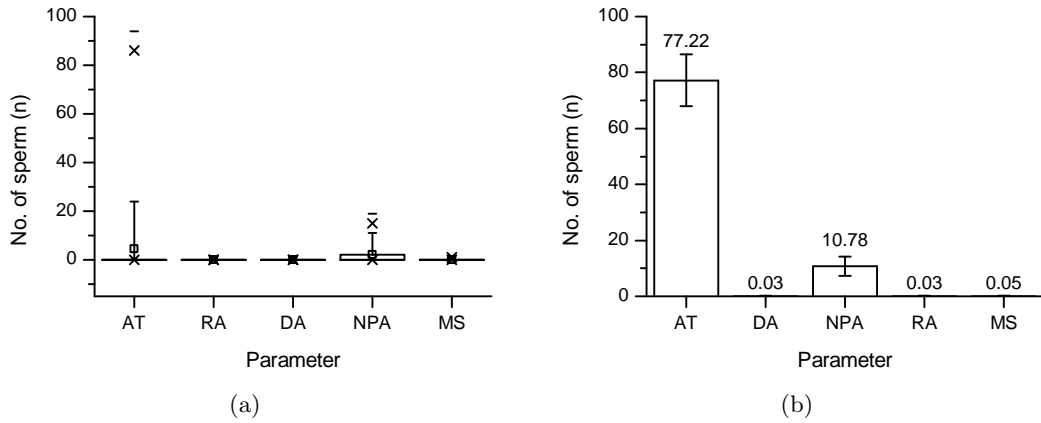


Figure 7.5: (a) The range and distribution of individual measurements when each parameter is varied through its full range and (b) the difference in the sperm population in the ampulla after 8 hours which can be obtained by varying individual parameters through their entire range. The error bars show S.D. AT = [Progressive Attachment Threshold], DA = [Max. Detachment Angle], NPA = [Max. Non-Progressive Angle], RA = [Max. Reflection Angle], MS = [No. of Movement Steps].

The [Progressive Attachment Threshold] was identified as the most significant parameter, with a maximum potential number of sperm in the ampulla of 94. How the number of sperm in the ampulla changes when the [Progressive Attachment Threshold] is varied through its full range is shown in Figure 7.6(a). This chart indicates that the [Progressive Attachment Threshold] must be between 0 and 0.3, as larger values result in no sperm progressing to the site of fertilisation. Further simulations were run to investigate the range 0 to 0.3 in 0.01 step intervals, as shown in Figure 7.6(b). This parameter is clearly highly sensitive to change.

With no quantitative literature to guide the setting of this parameter, the chosen value was estimated based on the curve shown in Figure 7.6(b). As described in Section 7.1.4, the pooled mean average for the number of sperm at the site of fertilisation at 9 hours post coitus from numerous studies (averaged from Table 7.1) is 21. Figure 7.6(b) shows a red horizontal line marking the point where the number of sperm is equal to 21. The corresponding [Progressive Attachment Threshold] at the point where the horizontal line crosses the curve is 0.1. Therefore, the estimated value for the [Progressive Attachment Threshold] was set to 0.1.

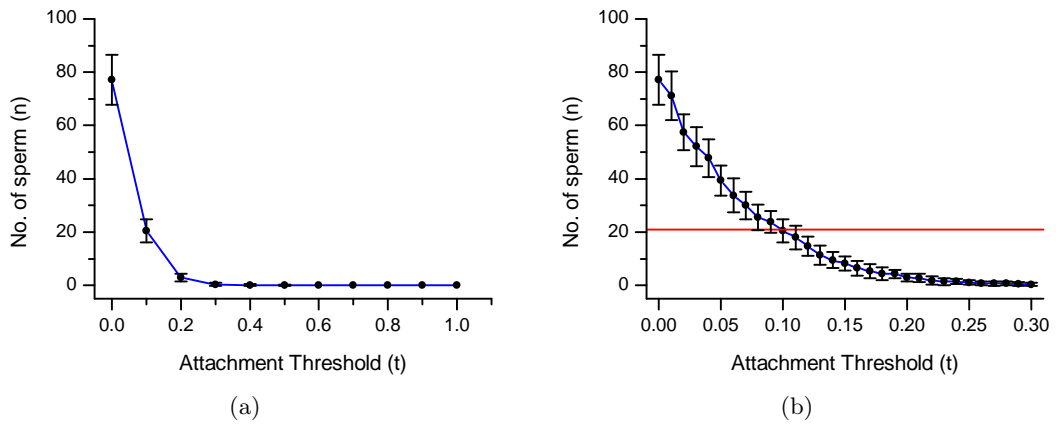


Figure 7.6: (a) The impact that varying the [Progressive Attachment Threshold] through its full range has on the number of sperm reaching the ampulla after 8 hours and (b) the impact of small changes over the identified subrange.

Now that a more realistic [Progressive Attachment Threshold] has been chosen, the remaining uncertain parameters were run again using the new value. The [Max. Reflection Angle], [Max. Detachment Angle] and [Max. Non-Progressive Angle] were varied between 0 and 180 degrees in 20 degree intervals. The [No. of Movement Steps] was varied between 1 and 12 in intervals of 1. The results are shown in the box plot in Figure 7.7. These results show that the [Max. Reflection Angle] has the largest potential impact after the attachment threshold, and the [Max. Non-Progressive Angle] has the second largest potential impact. The [Max. Detachment Angle] and [No. of Movement Steps] do not appear to have a large impact on the distribution of sperm.

The impact of each of these parameters when varied through their full range are shown in more detail in Figure 7.8. Figure 7.8(a) shows the impact that variations in the [Max. Detachment Angle] have on the final population in the ampulla. These show that there is a slight difference over the full parameter range, but that this difference is not significant. This indicates that this parameter is not very significant, so less

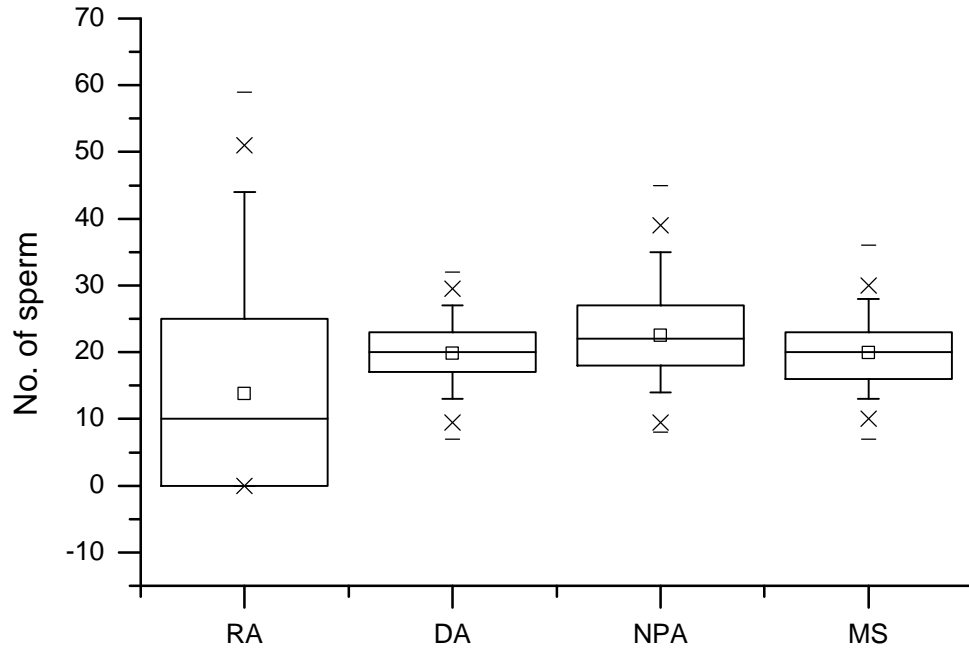


Figure 7.7: The difference in the sperm population in the ampulla after 8 hours that can be obtained by varying individual parameters through their entire range. The error bars show S.D. . DA = [Max. Detachment Angle], NPA = [Max. Non-Progressive Angle], RA = [Max. Reflection Angle], MS = [No. of Movement Steps]

accurate values are not going to influence the final result. The value of this parameter was kept at the initial value of 90 degrees.

Figure 7.8(b) shows the effect of varying the [No. of Movement Steps] has on the final distribution. As expected, the difference between 1 movement step and 2 movement steps has the largest difference, but after 2 the difference becomes negligible. Therefore the initial value of 4 was determined to be sufficiently accurate.

Figure 7.8(c) shows the effect of varying the [Max. Non-Progressive Angle] through its complete range. This shows that once the angle becomes 60 degrees or greater, the impact of variations in the distribution becomes negligible. An angle of greater than 60 degrees makes it more likely that non-progressive sperm will to turn around and swim back in the direction it started from, thus meaning that the turning direction of non-progressive sperm is completely random.

Figure 7.8(d) shows the influence of varying the [Max. Reflection Angle] through

its complete range. Values of less than 80 are significantly different from one of more values above 80. As the angle becomes greater than 80, the number of sperm reaching the ampulla tends to 0. At an angle of 0, the sperm do not turn far enough away from the surface to move away, and the likelihood of reattachment is increased. An angle of around 20 is optimal, as the sperm will follow the contours of the oviduct, guiding it straight to the site of fertilisation. Based on the data in Table 7.1, a reasonable value for this parameter would be between 40 and 60. Further simulations were run to focus on the range 35 to 60 in 5 degree intervals. These results are shown in Figure 7.9. As the initially chosen value of 45 is within the relative range (40-60), and there is no way of accurately determining the correct value of this parameter, the initial value is kept. The results do support the assumption that the reflection angle should be below 90 degrees.

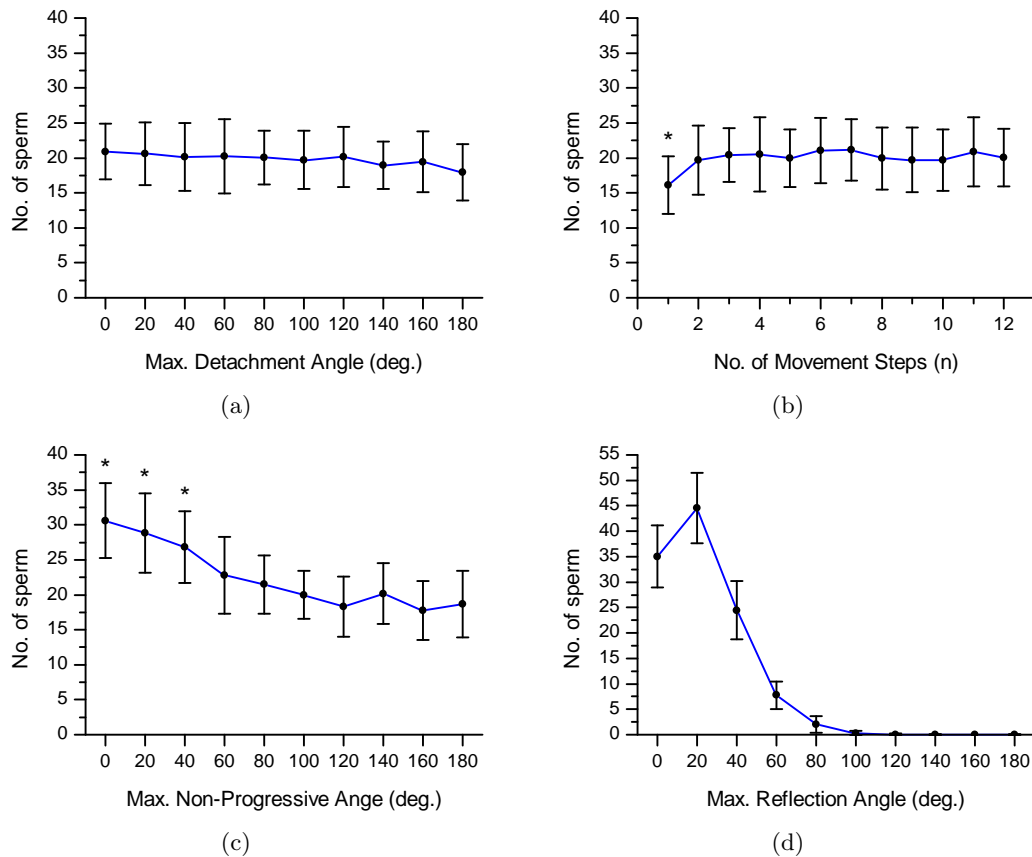


Figure 7.8: How variations of (a) the [Max. Detachment Angle] (DA), (b) the [No. of Movement Steps] (MS), (c) the [Max. Non-Progressive Angle] (NPA), and (d) the [Max. Reflection Angle] (RA) parameters influence the number of sperm in the ampulla after 8 hours.

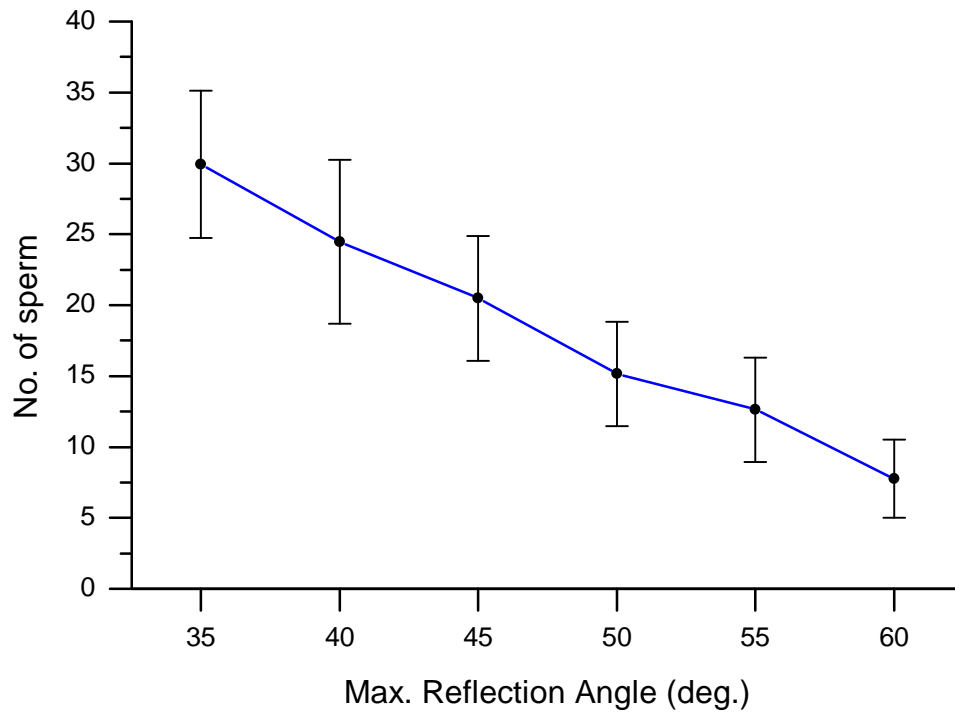


Figure 7.9: The impact of variations in the [Max. Reflection Angle] on the distribution of sperm within the range 35 - 60 degrees.

These simulations have resulted in a set of parameters which gives a similar number of sperm in the ampulla at the end of the simulation when compared to literature. The [Max. Detachment Angle] has been shown to have little or no impact, and the [No. of Movement Steps] is only significant at values lower than 2. The [Max. Reflection Angle] has been shown to be limited to 0-90, which supports the assumption made in Chapter 4.

### 7.2.3 Fully calibrated model

Table 7.2 shows the final set of system parameter values after calibration. The parameters which have been modified from their initial value are highlighted in bold.

The simulation was run 9 times, with 40 replicates per simulation run, generating a total of 360 replicates. Figure 7.10 shows the mean average distribution of sperm within the oviduct after 8 hours. The average number of sperm in the ampulla is not significantly different from the average of the data from Table 7.1.

The implemented model of sperm behaviour does not include some of the more complex

Parameter	Value	Source
Sperm Radius	$1.6 \mu m$	Measured (a)
Oocyte Radius	$36.65 \mu m$	Measured (b)
<b>Capacitation Threshold*</b>	<b><math>2.13675 \times 10^{-05}</math></b>	<b>Calibrated</b>
<b>Progressive Attachment Threshold*</b>	<b>0.1</b>	<b>Calibrated</b>
Non-Progressive Attachment Threshold	0.0	Observed (d)
Progressive Detachment Threshold	0.0	Observed (d)
Non-Progressive Detachment Threshold	1.0	Observed (d)
Time in Progressive State	$10s - 60s$	Observed (d)
Time in Non-Progressive State	$5s - 10s$	Observed (d)
Capacitated Sperm Life	$7200s$	Anecdotal (e)
Oocyte Maturation Period	$9000s$	Observed (f & g)
Progressive Velocity	$146.9 \mu ms^{-1}$	Measured (h)
Non-Progressive Velocity	$73.3 \mu ms^{-1}$	Measured (h)
No. of Movement Steps*	4	Estimated (h)
Max. Reflection Angle*	$45^\circ$	Estimated (d)
Max. Non-Progressive Angle*	$90^\circ$	Estimated
Max. Detachment Angle*	$90^\circ$	Estimated (d)
Block to Polyspermy Time	$2400s$	Observed (i)

Table 7.2: Final values of system level parameters used in the model after model calibration. The parameters with estimated or uncertain values are marked with an asterisk. (a) Firman and Simmons (2010), (b) Kim et al. (2004), (c) McGaughey et al. (1968), (d) Demott and Suarez (1992), (e) Bedford (1983), (f) Zamboni (1972), (g) Braden and Austin (1954a), (h) Goodson et al. (2011), (i) Wolf (1978)

behaviours described in literature, such as thermotaxis, chemotaxis and fluid dynamics. Based on the simulation results, if it is assumed that the model is completely accurate, then it appears that these complex behaviours do not have a significant impact on the reproductive process in the mouse. However, due to the presence of several uncertain parameters within the system, it is possible that the more complex mechanisms are being compensated for by the calibrated parameters.

The calibrated value for the [Progressive Attachment Threshold] of 0.1 is very low when compared to the value of 0.5 which was estimated from the qualitative description in literature. The qualitative description that this value was based on, “Sometimes the sperm would reattach the first time it contacted the epithelium, and sometimes it would bounce off, frequently making a direction change, and then continue swimming freely” (Demott and Suarez, 1992), is very vague, so a difference is not surprising. However, as the difference is so large, it is possible that this parameter is compensating for mechanisms which are not present in the model.

*In vitro*, human sperm have been observed to turn as they approach the wall of a

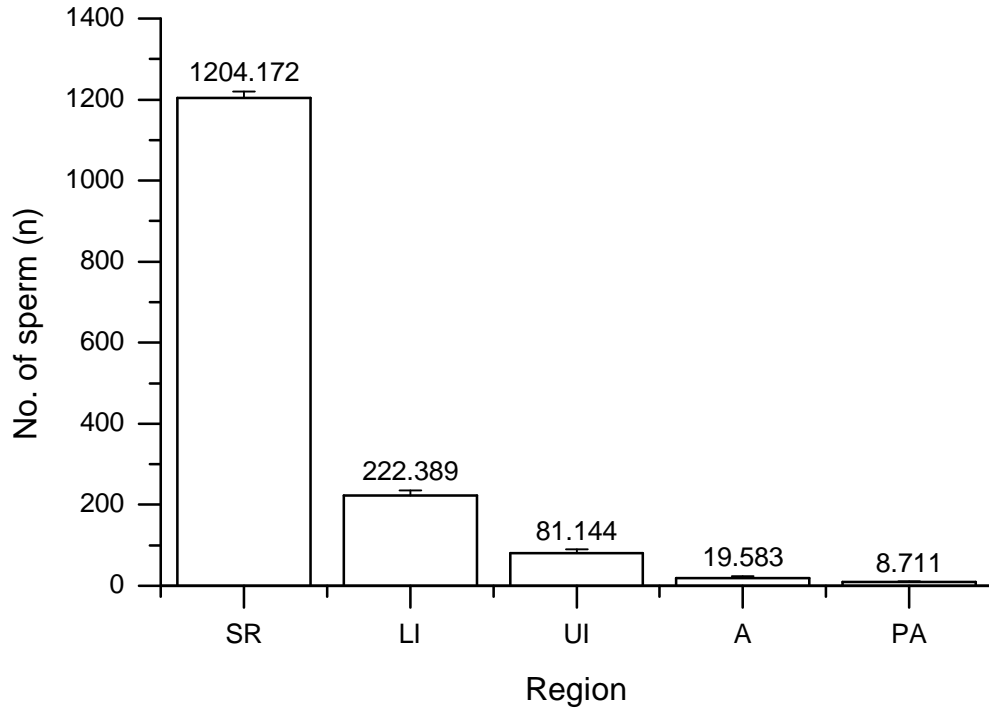


Figure 7.10: Final distribution of sperm within the oviduct after 8 hours using the calibrated parameters. SR = sperm reservoir, LI = lower isthmus, UI = upper isthmus, A = ampulla, PA = pre-ampulla.

chamber, and orient themselves to swim parallel to that wall (Rothschild, 1963; Smith et al., 2009; Woolley, 2003). If this behaviour also exists *in vivo* in mice, then the number of contacts between the sperm and the oviduct epithelium would be reduced. In addition to this, if sperm are partially carried in moving fluid due to peristaltic contractions, as has been previously hypothesised (Brokaw, 2001; Fauci and Dillon, 2006; Riffell and Zimmer, 2007), this could help to keep sperm away from the oviductal wall, thus reducing the opportunity for an attachment to be made.

The [Max. Reflection Angle] controls the progress of sperm along the oviduct. When the angle is low, sperm closely follow the oviduct wall, resulting in rapid progression of sperm to the site of fertilisation. Due to the fact that sperm move forward during progressive movement and only slightly deviate from a straight path. It is reasonable to assume that the angle that sperm reflect from the oviduct is narrow. Gaining a more accurate value for this parameter would improve the reliability of the model.

The [Max. Detachment Angle] does not appear to have a large influence on the distribution of sperm. As the sperm detach while in non-progressive movement, it is likely



that they may swim off in a random direction. Unlike sperm which are reflecting off the oviduct wall without attachment, sperm which are detaching from the oviduct are in a non-progressive movement state. Due to the large, erratic movements which are characteristic of non-progressive sperm, they may well change and swim off at a wider angle than sperm which do not attach. However, the same space restrictions to movement within the oviduct still apply, so they may not be able to fully turn around when detaching. Therefore, the chosen potential range of 0 to 90 is an acceptable value for the parameter.

The [Max. Non-Progressive Angle] is more difficult to characterise. When in a non-progressive movement state, some sperm swim erratically, with no identifiable path and others move in a noticeable direction, with occasional random deviations from the path. When the value of [Max. Non-Progressive Angle] is high, then the sperm can potentially swim off in any direction. When the value is low, then the sperm swimming direction is more constrained. However, the analysis of this parameter has shown that once the maximum potential angle becomes greater than 60, then the parameter no longer has a significant impact on the distribution of sperm. One way the model could be improved in the future is to have different populations of sperm, some which move in an identifiable direction when non-progressive, and others which can move completely randomly.

Combinations of the angle parameters could be compensating for both chemotaxis and thermotaxis. The simulation constrains the value of these parameters to within a specific range. The angle of rotation of sperm in different states could also be constrained *in vivo* due to temperature changes or the chemical composition of the oviduct.

Once a reasonable set of parameters had been established, the sensitivity of all parameters to change was investigated.

### 7.3 Sensitivity analysis

This section describes the sensitivity analysis performed on the system.

#### 7.3.1 Sensitivity to system parameters

The sensitivity of the system to variations for all system parameters was investigated. The value of each parameter was varied between -10% and +10% and the results are shown in Figures 7.11 and 7.12. Significance is determined using a wilcoxon signed-rank test when compared to the calibrated data set. The results show that using the fully calibrated values, the three parameters which are significantly sensitive to change

are the [Max. Reflection Angle], [Progressive Attachment Threshold] and the [Capacitated Sperm Life]. Figures 7.13 and 7.14 show the impact of sensitivity analysis on the number of penetrated oocytes. Only the [Progressive Attachment Threshold] is significantly different.

These results show that the [Progressive Attachment Threshold], [Max.-Reflection Angle] and the [Capacitated Sperm Life] are highly sensitive to change, and any hypotheses investigated using the model should take this into consideration.

### 7.3.2 Sensitivity to morphological parameters

One of the limitations of the 3D environments which was identified in Chapter 5 is that the histology process results in uneven shrinkage of the tissue by approximately 10%. In order to investigate the influence that this may have on the simulation, the scale of the lumen of the 3D environment was reduced by applying a scaling factor to the individual cross sections before binding them to the curve. The scale of the binding curve and the transformation matrix for each cross section was unchanged. This resulted in a reduction of the swimming space in the lumen without reducing the length of the tube. Scaling was applied to represent 10% shrinkage and 20% shrinkage, and a set of simulations was run for the resulting models. The results are shown in Figure 7.15.

Figures 7.15(a) to 7.15(e) show the impact that luminal shrinkage has on the distribution of sperm within the oviduct. These results do not show the entire population of sperm, just the sperm which have become capacitated during the simulation, which provides a better comparison for sperm distribution in the sperm reservoir. The results show that with 10% shrinkage, there is a significant increase in the number of sperm in the sperm reservoir, and a significant decrease in the number of sperm in the lower isthmus. However, there is no significant difference in the number of sperm in the upper isthmus, the ampulla or the pre-ampulla. This means that although fewer sperm are able to leave the sperm reservoir, those that do are more likely to progress to the upper part of the reproductive tract.

When the shrinkage is increased to 20% reduction, the number of sperm able to leave the sperm reservoir is significantly reduced, significantly reducing the number of sperm in the lower and upper isthmus and therefore significantly reducing the number of sperm reaching the site of fertilisation.

Figure 7.15(f) shows the impact that luminal shrinkage has on the number of sperm

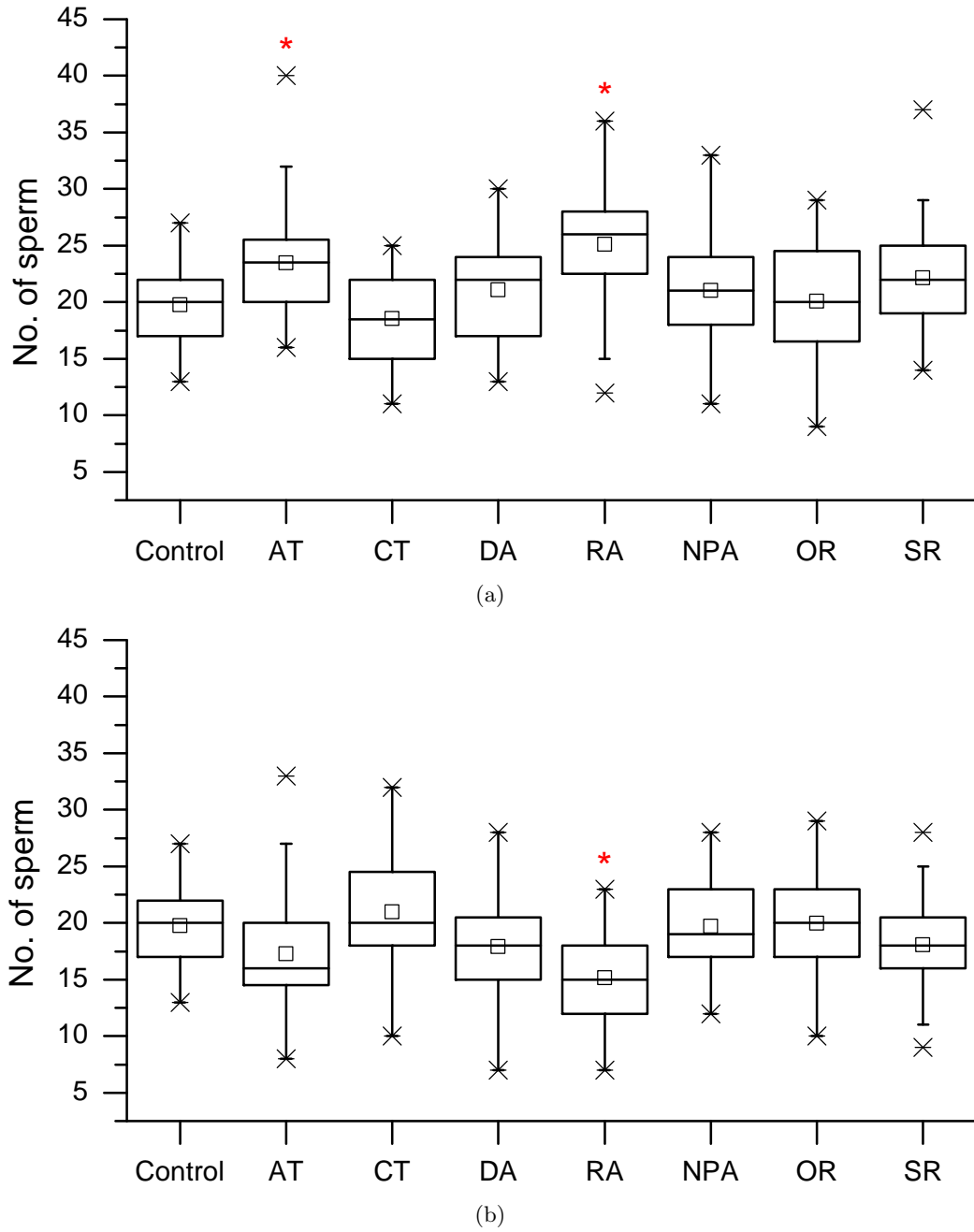


Figure 7.11: Results from sensitivity analysis of all system parameters in relation to the number of sperm in the ampulla after 8 hours, with (a) values reduced by 10% and (b) values increased by 10%. Control = The calibrated average, AT = [Progressive Attachment Threshold], CT = [Capacitation Threshold], DA = [Max. Detachment Angle], RA = [Max. Reflection Angle], NPA = [Max. Non-Progressive Angle], OR = [Oocyte Radius], SR = [Sperm Radius].

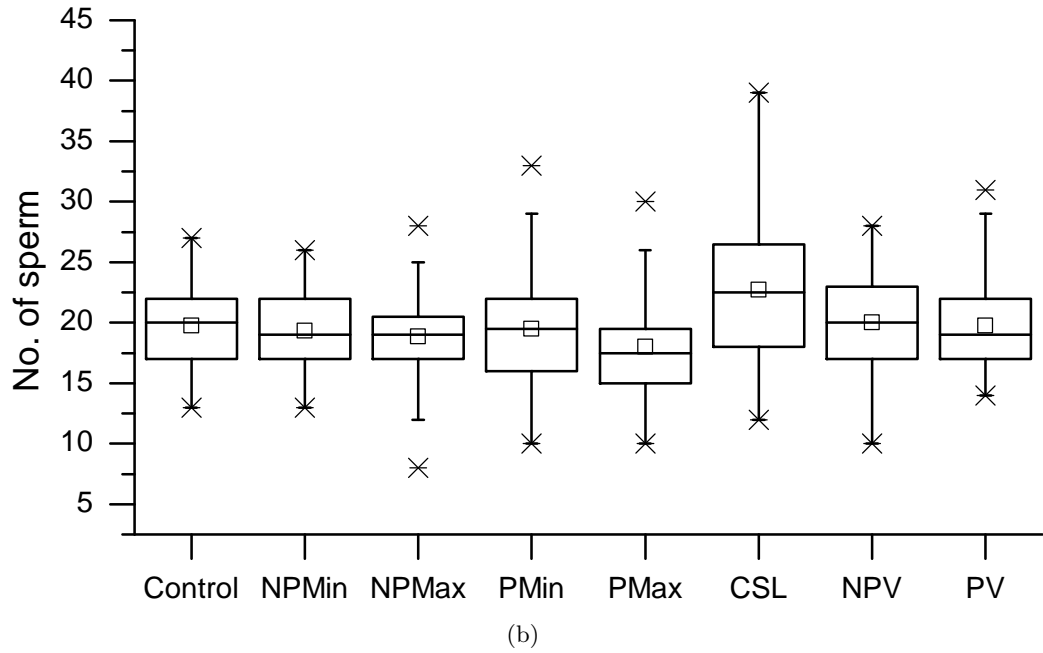
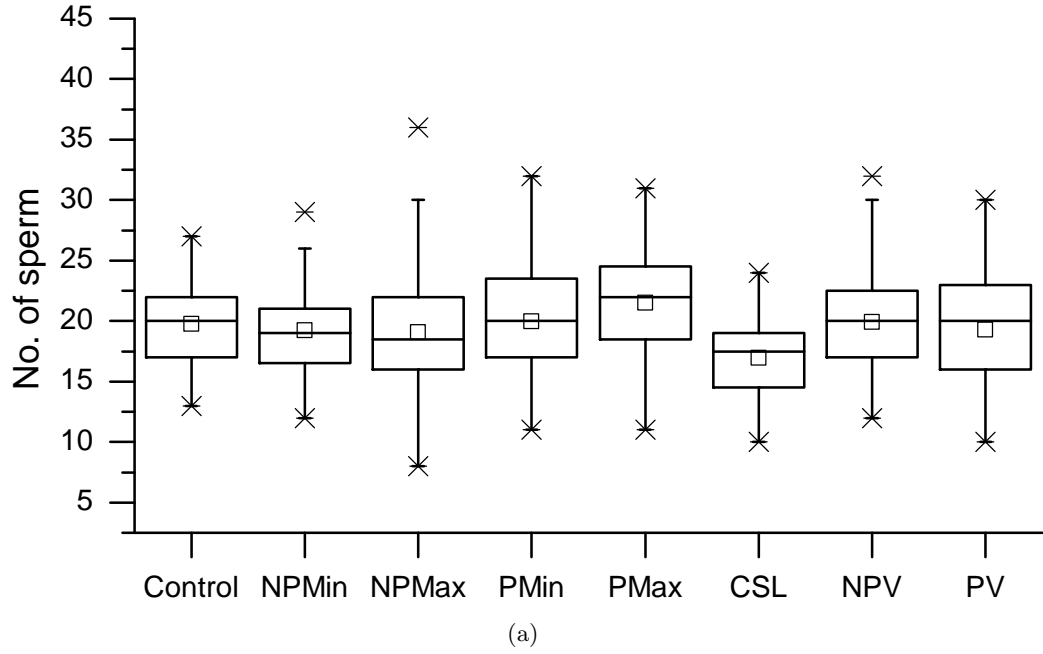


Figure 7.12: Results from sensitivity analysis of all system parameters in relation to the number of sperm in the ampulla after 8 hours, with (a) values reduced by 10% and (b) values increased by 10%. Control = The calibrated average, NPMIn = lower boundary of the [Time in Non-Progressive State], NPMMax = upper boundary of the [Time in Non-Progressive State], PMin = lower boundary of the [Time in Progressive State], PMax = upper boundary of the [Time in Progressive State], CSL = [Capacitated Sperm Life], NPV = [Non-Progressive Velocity], PV = [Progressive Velocity].

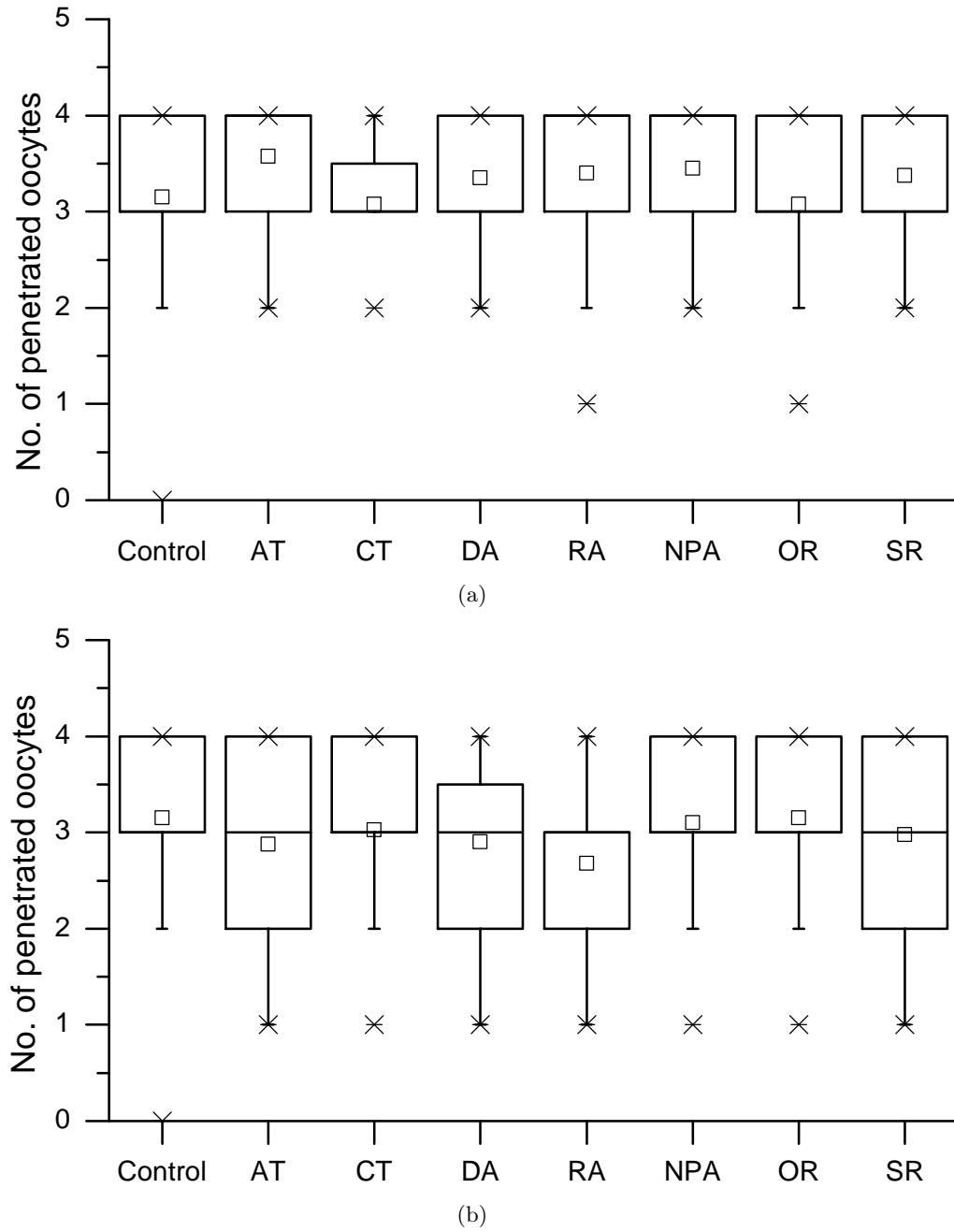


Figure 7.13: Results from sensitivity analysis of all system parameters in relation to the number of penetrated oocytes, with (a) values reduced by 10% and (b) values increased by 10%. Control = The calibrated average, AT = [Progressive Attachment Threshold], CT = [Capacitation Threshold], DA = [Max. Detachment Angle], RA = [Max. Reflection Angle], NPA = [Max. Non-Progressive Angle], OR = [Oocyte Radius], SR = [Sperm Radius].

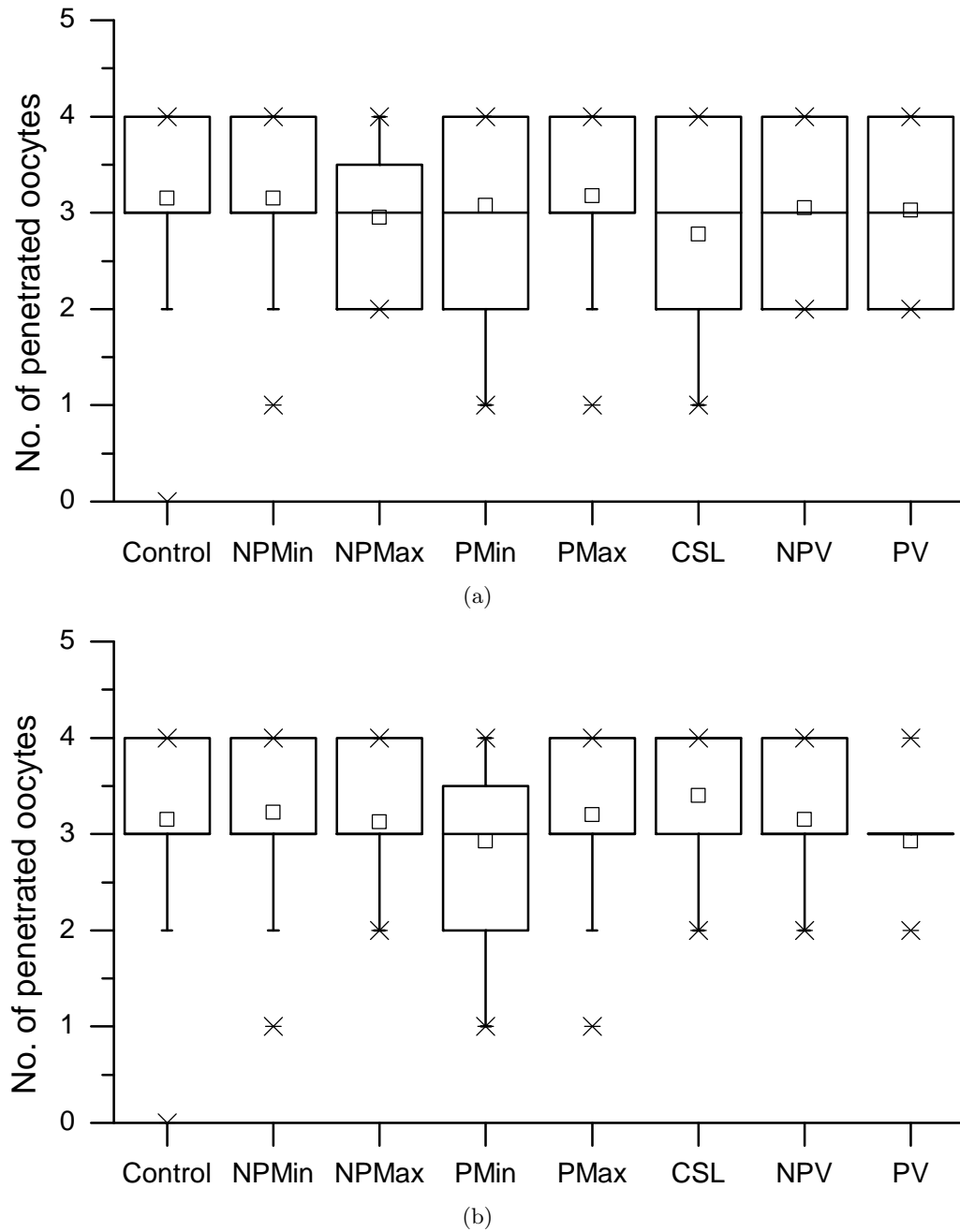


Figure 7.14: Results from sensitivity analysis of all system parameters in relation to the number of penetrated oocytes, with (a) values reduced by 10% and (b) values increased by 10%. Control = The calibrated average, NPMin = lower boundary of the [Time in Non-Progressive State], NPMMax = upper boundary of the [Time in Non-Progressive State], PMin = lower boundary of the [Time in Progressive State], PMax = upper boundary of the [Time in Progressive State], CSL = [Capacitated Sperm Life], NPV = [Non-Progressive Velocity], PV = [Progressive Velocity].

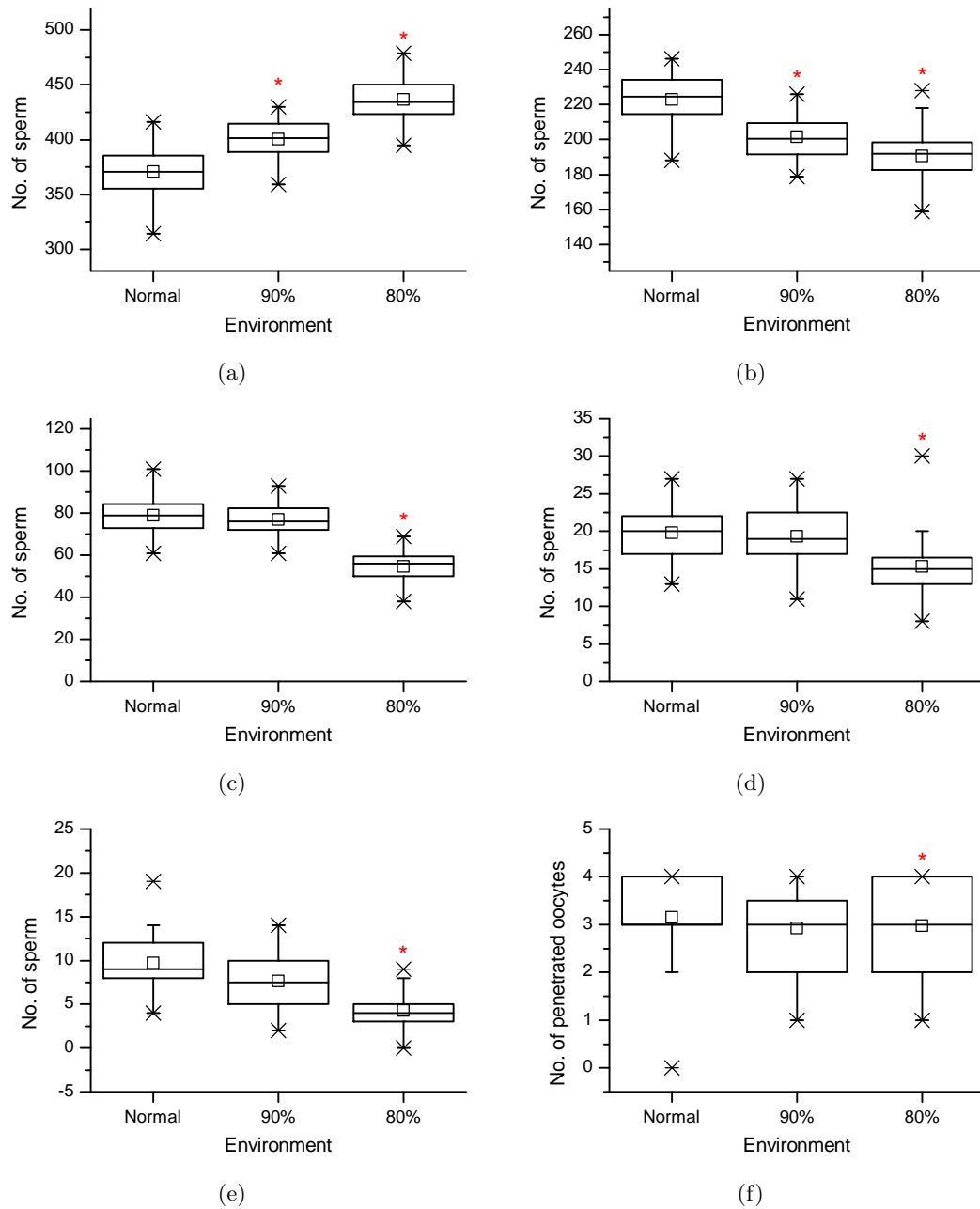


Figure 7.15: The impact of shrinking the lumen to 90% and 80% or its original size when compared to the Normal environment size on the number of sperm in (a) the sperm reservoir, (b) the lower isthmus, (c) the upper isthmus, (d) the ampulla, (e) the pre-ampulla and (f) the number of penetrated oocytes.

meeting the oocyte. Regardless of the amount of shrinkage applied, there is no significant difference in oocyte penetration.

Table 7.3 shows the relative number of sperm in each region as a percentage of the previous region. Regardless of the tissue shrinkage applied, a similar percentage of upper isthmus sperm are in the ampulla.

<b>% of Group</b>	<b>Normal</b>	<b>10%</b>	<b>20%</b>
SR in LI	59.61	50.33	43.61
LI in UI	36.49	38.13	28.76
UI in A	24.13	25.09	27.93
A in PA	44.48	39.43	28.10

*Table 7.3:* How the relative distribution of sperm changes as the shrinkage is increased. SR = sperm reservoir, LI = lower isthmus, UI = upper isthmus, A = ampulla, PA = pre-ampulla.

These results show applying a 10% reduction to the luminal swimming space does alter the distribution of sperm, but it does not have a significant impact on the number of reaching the ampulla. Applying a 20% reduction to the luminal swimming space does have a significant impact on the number of sperm reaching the ampulla, but it does not have a significant impact on the number of penetrated oocytes. From Table 7.3, with a 20% reduction applied, a smaller percentage of sperm from the lower isthmus move to the upper isthmus. However, a larger percentage of sperm from the upper isthmus are retained in the ampulla.

The implication of these findings are that a reduction in the lumen of the oviduct is beneficial to the system. Despite a reduction in the number of sperm reaching the site of fertilisation, the fact that the number of penetrated oocytes is the same means that the system is more efficient, and is likely to have fewer polyspermic oocytes.

## 7.4 Results validation against literature

This section describes the validation of the simulation results against literature. The model was calibrated against studies which reported the number of sperm in the ampulla at the end of the simulation. For validation, first the distribution of sperm and average number of capacitated sperm in the site of fertilisation were compared with different literature sources from those used for calibration. Then the percentage of polyspermic oocytes is compared with the reported number of polyspermic oocytes over time. Finally, two sets of simulations are setup to validate the number of penetrated oocytes over time with initial conditions matching the described experiments in two different studies. In total, seven different literature sources were used for validation.



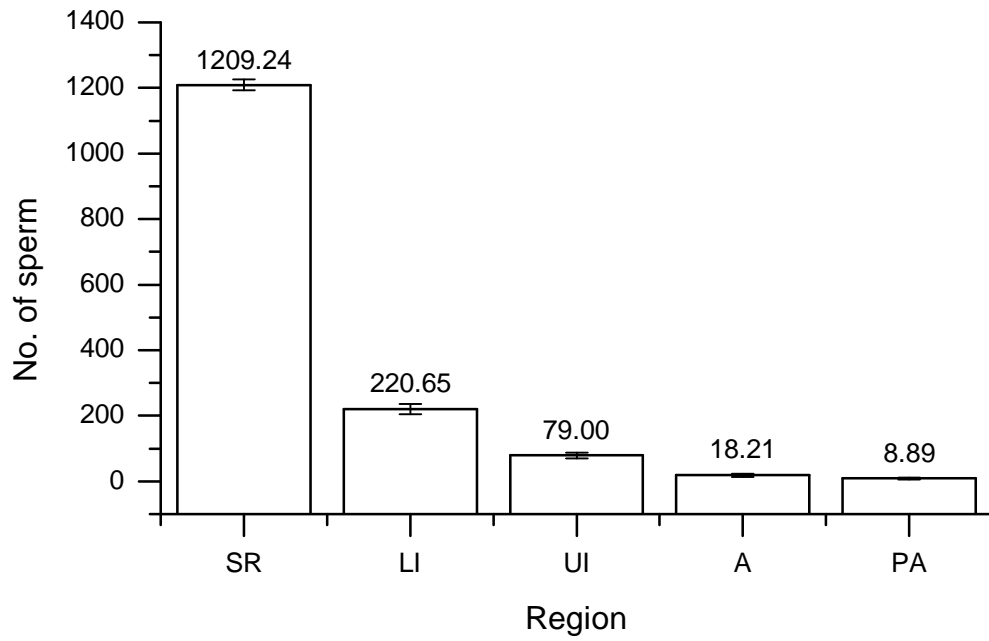
### 7.4.1 Comparison with literature

Up to this point, only a single 3D environment has been used for all simulations. However, in nature, the internal structure and external bends of each oviduct is different. In order to provide a more realistic dataset for validation, the simulations in this section were run using all 9 environments described in Section 5.7. These environments contain a cross of three different curve structure (1, 2 & 3) and three different internal fold structures (A, B & C). The pooled mean of all environments is used as a baseline for comparison.

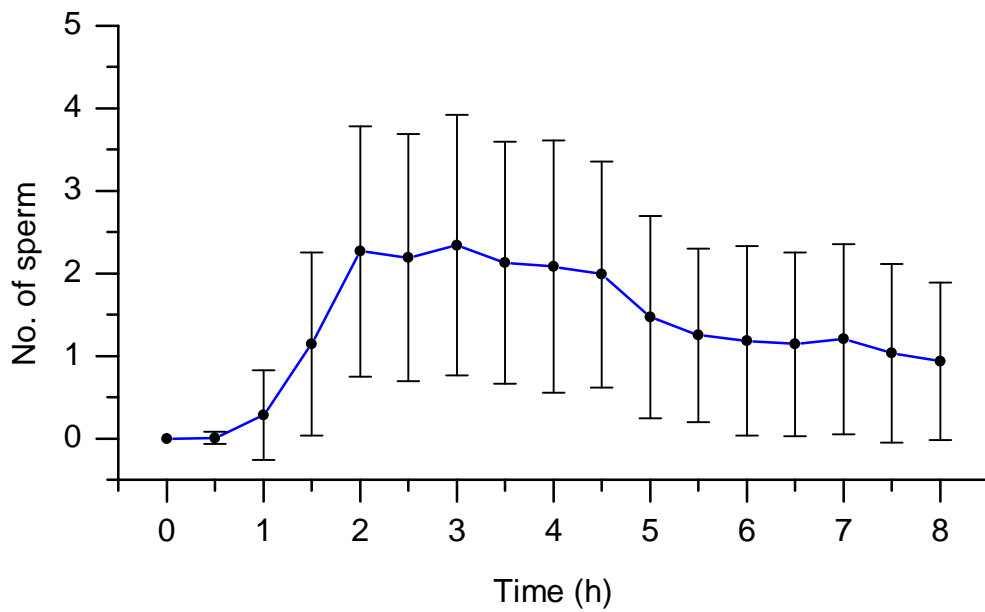
The pooled distribution of sperm at the end of the simulation from all environments are shown in Figure 7.16(a). When looking at the relative distribution of sperm, it has been shown in rabbits that 82%-99% of sperm were still in the first two segments of the oviduct after several hours, with between 1% to 5% in the ampulla (Harper, 1973). In the golden hamster, the percentage of sperm in the lower segments of the oviduct is 98%, with less than 1% in the ampulla (Smith et al., 1987). When looked at as a percentage of the total population, the simulation results show that 79% of the sperm are still in the sperm reservoir and 14% are in the lower isthmus. This means the total percentage of sperm in the lower segments of the oviduct in the simulation is between 79%-93%. The percentage of sperm in the ampulla in the simulation is only 1%. This demonstrates that the distribution of sperm is within a similar range to that observed in different species. Descriptions of the processes in reproductive biology are commonly based on evidence from different species, and the relative distribution of sperm in different mammals is similar (Eisenbach and Giojalas, 2006; Ikawa et al., 2010; Suarez, 2008b). However, the species' used for comparison are different from the modelled species, so these comparisons should be used primarily as an indicator of correctness, not as proof of validity.

Figure 7.16(b) shows the average number of capacitated sperm in the ampulla over time. It has been observed that only 1 or 2 capacitated sperm are seen in the ampulla of the mouse at any time (Zamboni, 1972). Tessler and Olds-Clarke (1981) reported that on average between  $0.7 \pm 0.3SEM$  and  $1.8 \pm 1.1SEM$  sperm were found free swimming in the mouse ampulla at different time points. When compared statistically, there is no significance between the number of capacitated sperm in the ampulla in the simulation and that reported in literature.

Polyspermy, which is when multiple sperm penetrate an oocyte, was estimated from the results as described in Section 4.3.2. The results are shown in Figure 7.17, along with a comparison to the percentage of polyspermic oocytes in the mouse reported in



(a)



(b)

Figure 7.16: (a) The pooled average distribution of sperm for all environments after 8 hours using the calibrated parameters and (b) the average number of capacitated sperm in the ampulla for all environments over time. SR = sperm reservoir, LI = lower isthmus, UI = upper isthmus, A = ampulla, PA = pre-ampulla.

two sets of data in Braden et al. (1954). The first source, ‘Lit. Table 1’ is from Table 1 (pg. 394) the second source ‘Lit. Table 2’ is from Table 2 (pg. 395). These results were compared using the Chi square method, and neither literature source was found to be significantly different from the simulation at  $p=0.001$ . This demonstrates that the technique applied for estimating incidence of polyspermy is reasonable. In order to provide a more comprehensive comparison with literature, two different studies were used and a set of virtual experiments were carried out, including the initial conditions of the corresponding studies.

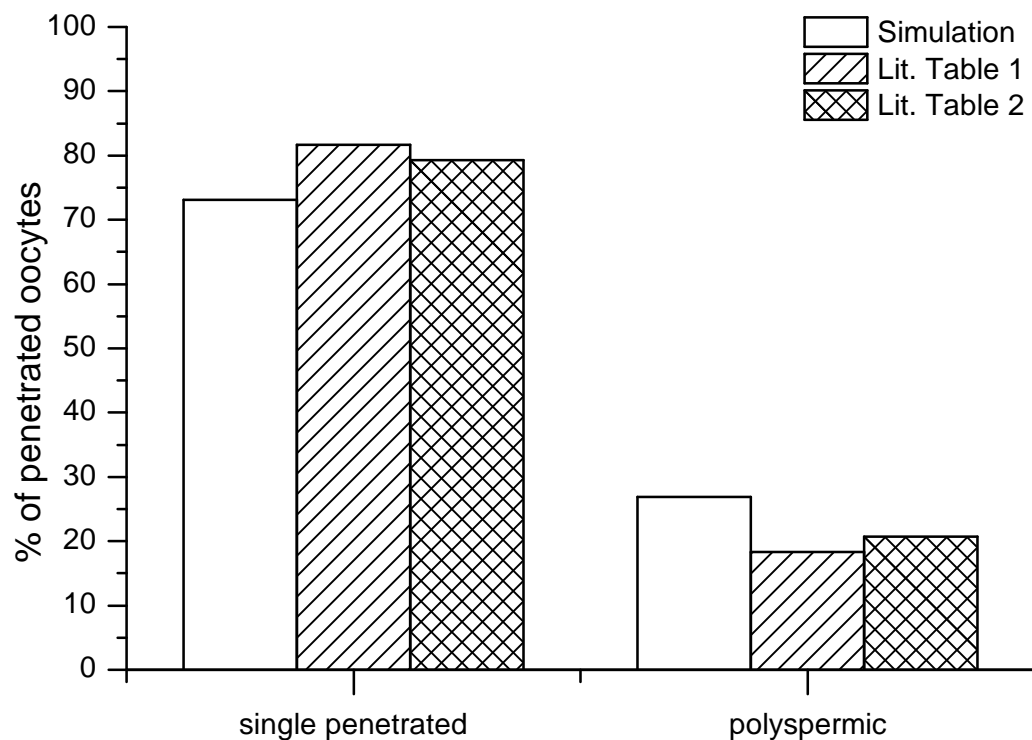


Figure 7.17: A comparison between the simulation results and the number of penetrated sperm which are polyspermic from Braden et al. (1954). ‘Lit. Table 1’ is from Table 1 (pg. 394) and ‘Lit. Table 2’ is from Table 2 (pg. 395).

#### 7.4.2 Validation study 1

The first study used for validation is Braden and Austin (1954a). The paper describes a set of experiments which aim to investigate the effect of delayed coitus and heat shock treatment on fertilisation in the mouse. Several experiments are presented. The experiment used for validation looks at the number of penetrated eggs in mice sacrificed at regular intervals after mating. The total number of penetrated eggs is presented, along with the number of mice which have penetrated eggs and the number of mice

where greater than 75% of the eggs are penetrated. In the experiments, all matings occur after ovulation, so in all mice oocytes can be fertilised as soon as the sperm reach them. A set of simulations was setup to compare the results of the model with the results of the study.

As biological systems naturally have high degrees of variation, it is not reasonable to expect the results to match exactly. Therefore, the hypothesis being tested is: The percentage of penetrated oocytes with respect to time in the model is significantly different from the biological system for more than 50% of the simulated time period.

The time of ovulation and the oocyte maturation period were both set to 0. All nine environments were used, and the results were pooled, giving 360 replicates in total. Figure 7.18 shows the percentage of eggs penetrated over time in both literature and in the simulation.

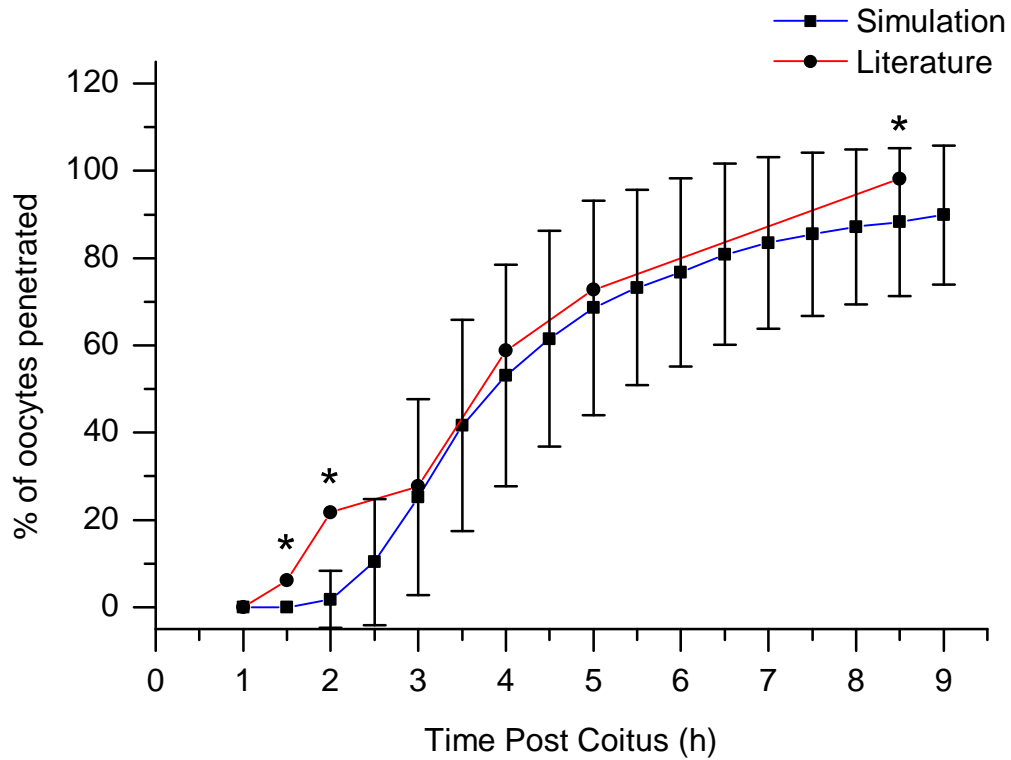


Figure 7.18: A comparison between the simulation results and data presented by Braden and Austin (1954a).

Out of the 6 time-points available the literature, three are significantly different from the simulation. However, both the simulation and literature follow the same path. There appears to be are two key differences between the results. First, the simulation appears

to be approximately 0.5 hours behind the real system and secondly, the simulation curve is not as steep as the curve in literature and the same percentage of fertilised eggs is not reached in the simulation.

There are several possible explanations for these discrepancies. The time plotted in Figure 7.18 is based on the reported time of killing in the study. However, this does not taken into account the fact that spermatozoa are motile individuals, and will still be able to move after the mouse has died, as confirmed by the *in situ* observations of sperm movement made by Suarez (1987), and could potentially penetrate the oocytes after the death of the animal. The study does not mention the amount of time that may have elapsed between killing the animal and collection of the eggs, although it is likely to be immediately after death. It is therefore reasonable to assume that all 10 mice could not be sacrificed, dissected and have their eggs isolated within the space of a few minutes, and it could potentially have been 10 - 15 minutes after the reported killing time.

The simulation was initialised to start 1 hour *post coitus*, with all sperm in the sperm reservoir. However in the real system, sperm is rapidly transported to the uterus after coitus, and have been observed in the oviduct within 15 minutes (Harper, 1982). As mating in the comparison study occurs after ovulation, as soon as sperm enter the oviduct, it is possible that a few sperm could become capacitated and ascend to the site of fertilisation in less than 1 hour. This means that in the real system, sperm could be closer to the eggs by the time the simulation starts.

In regard to the hypothesis, although oocyte penetration at some of the time points are statistically different, there is no statistical difference for the majority of the time points. Therefore, the hypothesis has be be rejected. This does not indicate that the model is an exact match for the biological system, but it does demonstrate that the model is able to predict oocyte penetration for the majority of the time points.

#### 7.4.2.1 Validation Study 1.1

The results generated by the simulation were also used for comparison with another study. Olds (1970) reported the average number of sperm at the site of fertilisation in mice to be  $6.67 \pm 7.31SD$  with a range of [0 - 25] when sperm are inseminated after ovulation. In the study, the average number of sperm was calculated based on measurements taken from 1 hour *post coitus* to 6 hours *post coitus*. This data was used to test the hypothesis: The average number of sperm over time in the ampulla in the model is significantly different from the biological system. When the number of sperm in the ampulla of the simulation is averaged over the same time period, the average

number is  $4.04 \pm 4.08SD$  with a range of [0 - 24]. There is no significant difference between the average reported in the study and the average in the simulation. Therefore, the hypothesis can be rejected. These results demonstrate that the rate at which sperm reach the site of fertilisation is a good match for the real system.

### 7.4.3 Validation study 2

In order to confirm the validity of the results, a second study by McGaughey et al. (1968) was used to validate the simulation. The paper describes a set of experiments to determine the maximum fertilising lifetime of sperm in mice. In the experiment used for validation, ovulation was induced in mice by HCG injection. Mice injected with HCG were previously observed to ovulate 10-12 hours after the injection (Marston and Chang, 1964). Subsets of mice were inseminated at regular intervals before the expected time of ovulation. This allows the relationship between the time of insemination and the time of ovulation to be investigated. Eggs were recovered 10 hours after ovulation and the number of fertilised eggs were counted.

A set of virtual experiments was setup to compare the results of the model with the results generated by this study. The hypotheses being tested are: (A) the percentage of penetrated oocytes in the model is significantly different from the biological system when insemination is at or around the time of ovulation and (B) The percentage of penetrated oocytes is significantly different when insemination is prior to ovulation. As the time of expected ovulation is within a 2 hour period, an estimated time matching the midpoint was used as a reference. The [Ovulation Time], which is an input to the model, was modified to match the expected time of ovulation in relation to the HCG injection.

A set of experiments was run for each 3D environment at each corresponding [Ovulation-Time]. The [Total Simulation Time] for each simulation was set to 10 hours after the [Ovulation Time]. How the timings reported in literature are mapped onto the [Ovulation Time] and the [Total Simulation Time] is shown in Table 7.4. The results were pooled for all environments and a comparison between the distribution in literature and the simulation is shown in Figure 7.19.

The results show that when sperm are inseminated within a few hours of ovulation, the number of penetrated oocytes is not significantly different to those in literature. However, when insemination occurs several hours before ovulation, the number of penetrated oocytes in the simulation is only slightly reduced, whereas the study reports a very small percentage of penetrated oocytes. This difference is possibly due to the fact that the study measures penetrated oocytes, whereas the model assumes that oocyte

Insemination after HCG (h)	Time prior to ovulation (h)	Ovulation Time (h)	Total Simulation Time (h)
-2	12 - 14	12	22
0	10 - 12	10	20
3	7 - 9	7	17
6	4 - 6	4	14
8	2 - 4	2	12
9	1 - 3	1	11
12	-2 - 0	-2	9

Table 7.4: How the timing of HCG injections reported in McGaughey et al. (1968) relates to the time of ovulation, the [Ovulation Time] input parameter and the [Total Simulation Time].

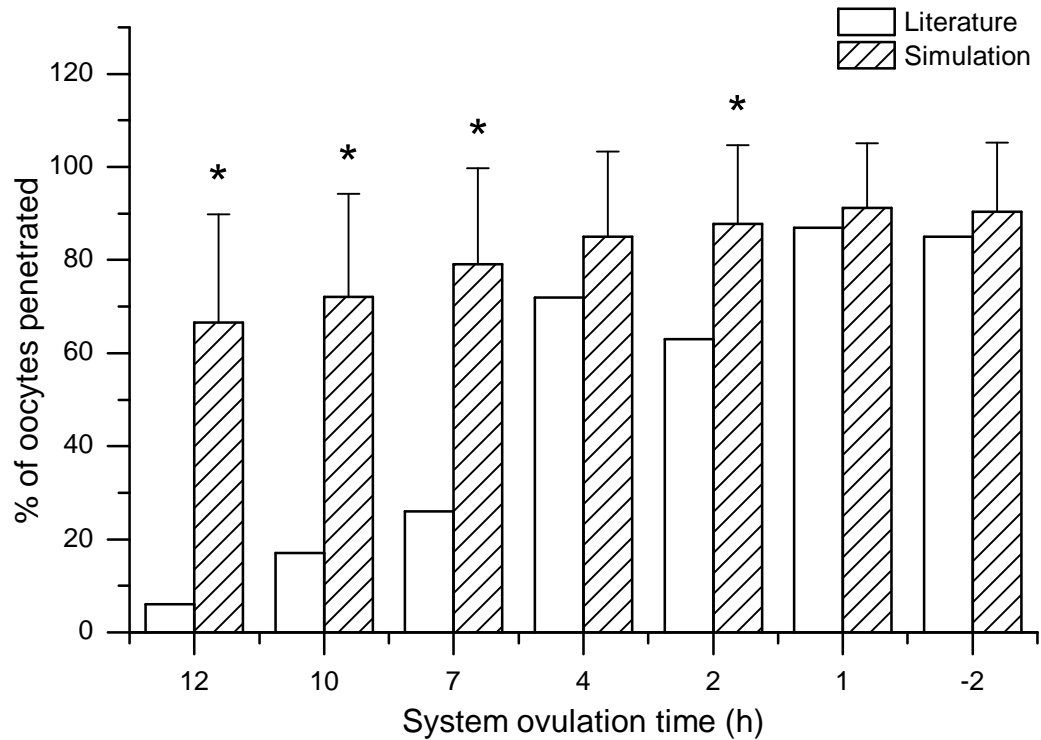


Figure 7.19: A comparison between the simulation results and data presented by McGaughey et al. (1968).

penetration occurs with every sperm attachment. It has been shown in mice that sperm which have been aged for a few hours before insemination are no longer able to penetrate the oocyte (Smith and Lodge, 1987).

With regard to hypothesis (A), although the penetrated oocytes at time points 4, 1 and -2 are not significantly different, the penetrated oocytes at time point 2 is significantly different. There is possibly an anomaly with the data from literature with the percentage of penetrated oocytes at 8 hours after the HCG injection, which matches the system ovulation time of 2 hours on Figure 7.19. The percentage of fertilised oocytes is 63%, which is significantly lower than the percentage at 9 hours after the HCG injection (87%) and 6 hours after the HCG injection (72%). When looking at the experimental design, the sperm used for insemination at 8 hours after the HCG injection were taken from the *vas deferens* and *cauda epididymidis*, which are regions of the male reproductive system, whereas the inseminations at 9 hours and 6 hours were taken from the uterus of a female mouse which had recently mated. Although it is convenient to assume that the literature source is incorrect, without a second literature source to confirm the anomaly, the hypothesis has to be accepted. With regard to hypothesis (B), the results clearly indicate that insemination prior to ovulation in the biological system is significantly different from the model, and therefore the hypothesis is accepted.

## 7.5 Conclusions

This chapter has demonstrated that the model is capable of predicting oocyte penetration over time with reasonable accuracy. The uncertain parameters ([**Progressive Attachment Threshold**] & [**Max. Reflection Angle**]) have been identified as highly significant, and neither has an accurate value which was derived experimentally. Therefore, the value of these parameters may be compensating for mechanisms which are not included in the model.

Within the model, the results with respect to the time of copulation in relation to ovulation have been validated. However, there are still several unknowns. The parameters which were calibrated or estimated need to be investigated in more detail for the model to truly be considered to be validated. Although the underlying mechanisms have not been fully validated, the model does generated results which are statistically similar to the real system. Therefore, the model can be used to test hypotheses about how variations in the system would influence successful sperm to oocyte interactions.

The [**Progressive Attachment Threshold**], [**Capacitated Sperm Life**] and the [**Max. Reflection Angle**] are significant factors with regard to the number of sperm



reaching the site of fertilisation, but only the [Progressive Attachment Threshold] is significant in relation to the number of penetrated oocytes.

A 10% reduction in the swimming space of the lumen to compensate for histology shrinkage does have a significant impact on the distribution of sperm, but not on the number of sperm reaching the ampulla. A 20% reduction in the swimming space of the lumen does have a significant impact the number of sperm reaching the ampulla. Neither 10% or 20% reduction in the swimming space of the lumen has a significant influence on the number of penetrated oocytes.

Different aspects of the system have been compared with different literature sources, and the distribution of sperm and number of capacitated sperm at the site of fertilisation at different time intervals are not statistically different from the values published in literature.

When incidence of polyspermy was calculated, the simulation is able to predict the percentage of oocytes with multiple penetrations. This indicates that the rate of sperm reaching the site of fertilisation is similar to the real system.

It has been shown that when compared to two different independent studies, the system is able to predict the same fertilisation rate when the coitus to ovulation timings are changed, but not when coitus is more than 4-6 hours before ovulation.

In Chapter 8, a set of experiments will be performed using the model to investigate how the 3D environment influences the reproductive process.

## Chapter 8

# Experimentation

The purpose of this chapter is to use the validated computational model of sperm behaviour to attempt to address the research questions posed in Chapter 1. First, a set of experiments to investigate the impact of variations in the 3D environments on sperm distribution, oocyte penetration and incidence of polyspermy were carried out. Secondly, a set of experiments to investigate the impact that variations in the number of sperm in the oviduct has on the distribution of sperm, the number of fertilised oocytes and incidence of polyspermy were conducted. The results are discussed in relation to the research questions posed in Chapter 1.

### 8.1 Experiment 1: modifying the 3D environment

The 3D structure of the oviduct has been identified as a potential regulator of sperm progression to the site of fertilisation (Suarez, 1987; Zamboni, 1972). Both the external bends and the internal soft tissue folds may contribute to the regulation of sperm passage to the site of fertilisation.

This section attempts to address the research questions posed in Chapter 1, (c) “does the complex 3D structure of the oviduct regulate the passage of sperm to the site of fertilisation?” and (d) “Which aspects of the 3D model are most significant with respect to the regulation of sperm progression to the site of fertilisation?”

To address research question (c), the following hypotheses were used: (A) Removing structural features of the oviduct does not significantly change the percentage of penetrated oocytes and (B) Removing structural features of the oviduct does not significantly change the percentage of polyspermic penetrated oocytes. To address research question (d), the following hypothesis was used: (C) Removing the internal folds of the oviduct has a significantly larger impact on the percentage of penetrated oocytes than removing the external bends. With regard to analysis of experimental results

for this chapter, statistical significance is determined using a non-parametric Wilcoxon signed-rank test with a p value of 0.001.

### 8.1.1 Natural variations between individuals

In nature, the internal folds and external bends of each oviduct is different. In order to investigate the impact that differences in the structure of the oviduct has on the behaviour of the simulation, a set of simulations were carried out using all 9 environments described in Chapter 5. These environments contain a cross of three different curve structures (1, 2 & 3) and three different internal fold structures (A, B & C). For each environment, a simulation with 40 replicates was run for 8 hours. The pooled mean of all environments is used as a baseline for comparison.

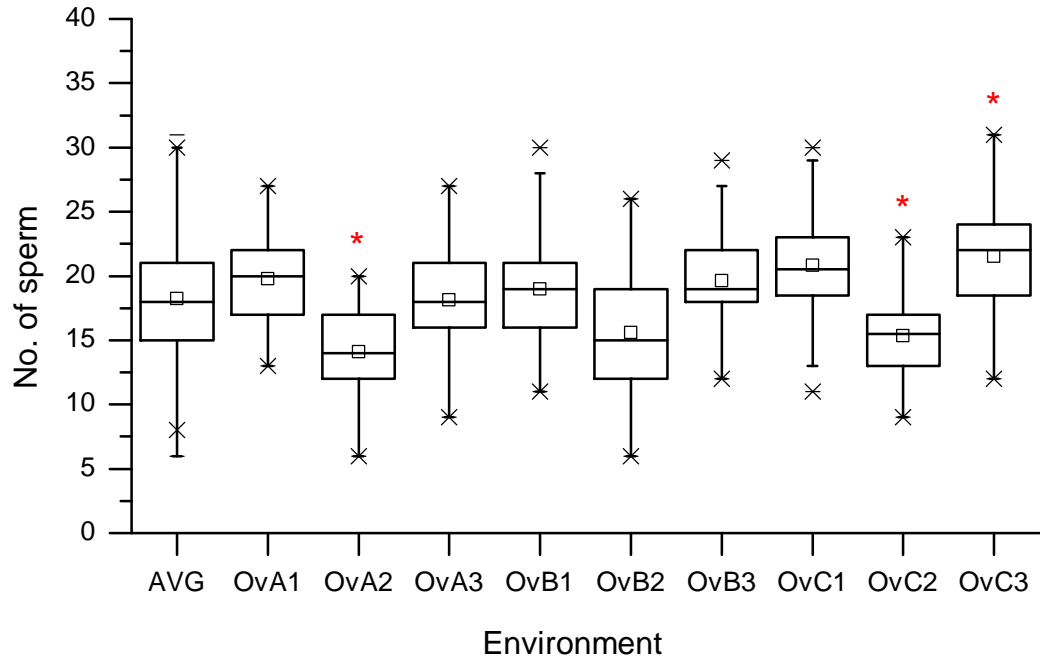
Figure 8.1(a) shows the impact that variations in the 3D environment have on the number of sperm in the ampulla. Several of the environments are significantly different from the pooled mean. When grouped by curve type (Figure 8.1(b)), a significant difference is observed in two of the curves. When grouped by fold type (Figure 8.1(c)), no significant difference is observed. This indicates that a change in curvature alone, but not a change in fold structure alone can have a significant impact on the progression of sperm. It also shows that the combination of both can have a significant impact.

When looking at the impact of changes in the 3D environment on the number of penetrated oocytes (Figure 8.2(a)) there is no significant difference from the pooled mean. When grouped by curve type (Figure 8.2(b)) and fold type (Figure 8.2(c)), no significant difference is shown. This shows that although differences in the 3D structure impact the distribution of sperm, they do not significantly influence with the number of sperm to oocyte encounters.

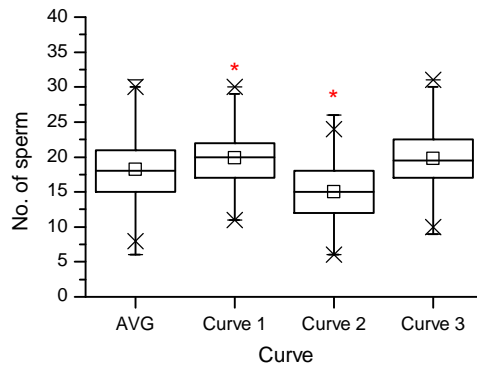
When looking at the impact on the percentage of penetrated oocytes which are polyspermic (Figure 8.3(a)) there is no significant difference from the pooled mean. When grouped by curve type (Figure 8.3(b)) and fold type (Figure 8.3(c)), no significant difference is shown. Therefore, variations in the environment do not alter the percentage of penetrated sperm which are polyspermic.

### 8.1.2 Relative significance of bends and folds

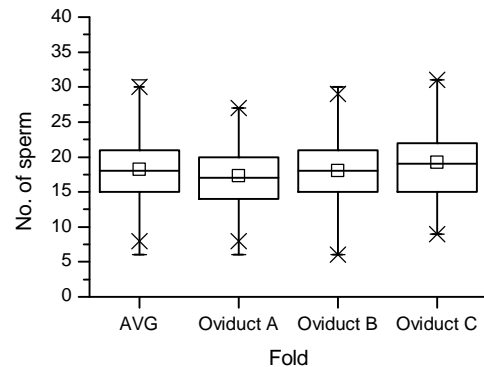
In order to determine the relative significance of the bends and the folds in the oviduct, and how the combination results in regulated passage of sperm to the site of fertilisation, two more sets of 3D oviduct models were constructed. The first set contained the three different 3D bend forms (1, 2 & 3) without any internal folds. The second set contained



(a)

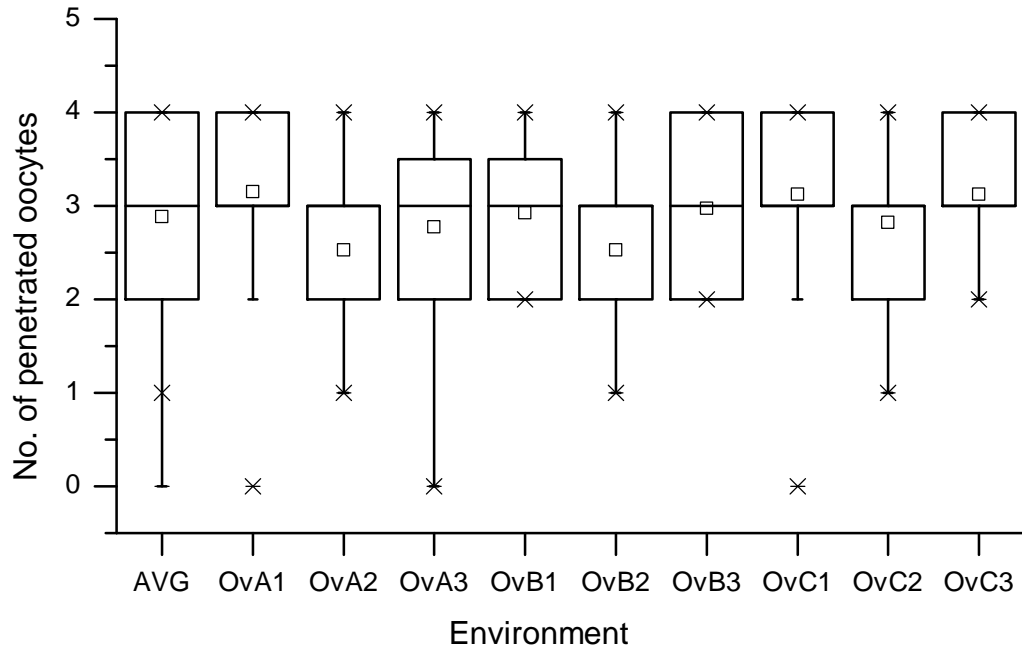


(b)

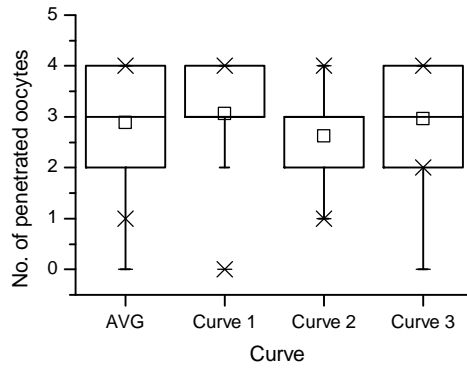


(c)

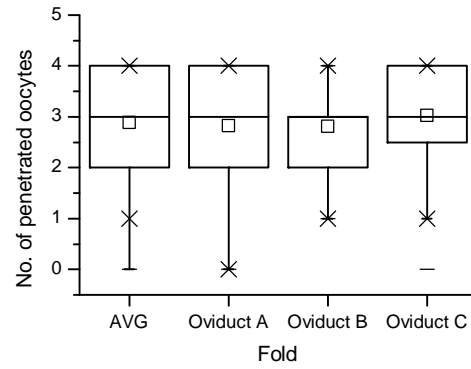
Figure 8.1: The impact that (a) individual variations in the 3D environments, (b) variations in the external bends and (c) variations in the internal folds have on the number of sperm in the ampulla.



(a)

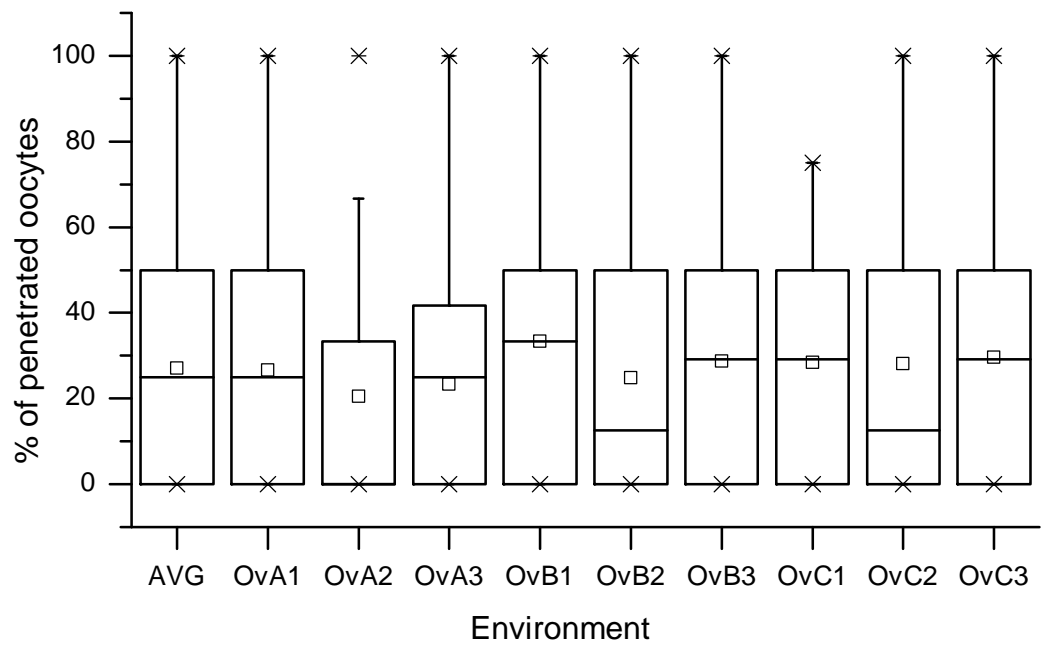


(b)

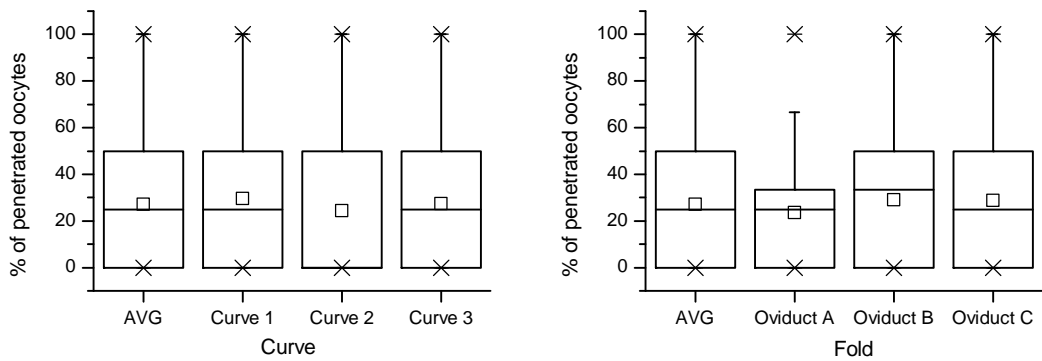


(c)

Figure 8.2: The impact that (a) individual variations in the 3D environments, (b) variations in the external bends and (c) variations in the internal folds have on the number of penetrated oocytes.



(a)



(b)

(c)

Figure 8.3: The impact that (a) individual variations in the 3D environments, (b) variations in the external bends and (c) variations in the internal folds have on the percentage of penetrated oocytes which are polyspermic.

the three different internal fold structures (A, B & C) without the external bends. A final model was constructed which contained no bends or folds. Where the models do not contain internal folds, the lumen is a cylindrical tube, with a radius equal to the Inner Radius measurements taken from the histology slides (see Section 5.4.3). Details of how the models were constructed are shown in Appendix C. Figure 8.4 shows how the relative distribution of capacitated sperm changes in different regions of the oviduct as the 3D structure is altered. Sperm which did not become capacitated during the simulation were not included. Excluding sperm which did not become capacitated gives a clearer indication of how the population of sperm in the sperm reservoir changes.

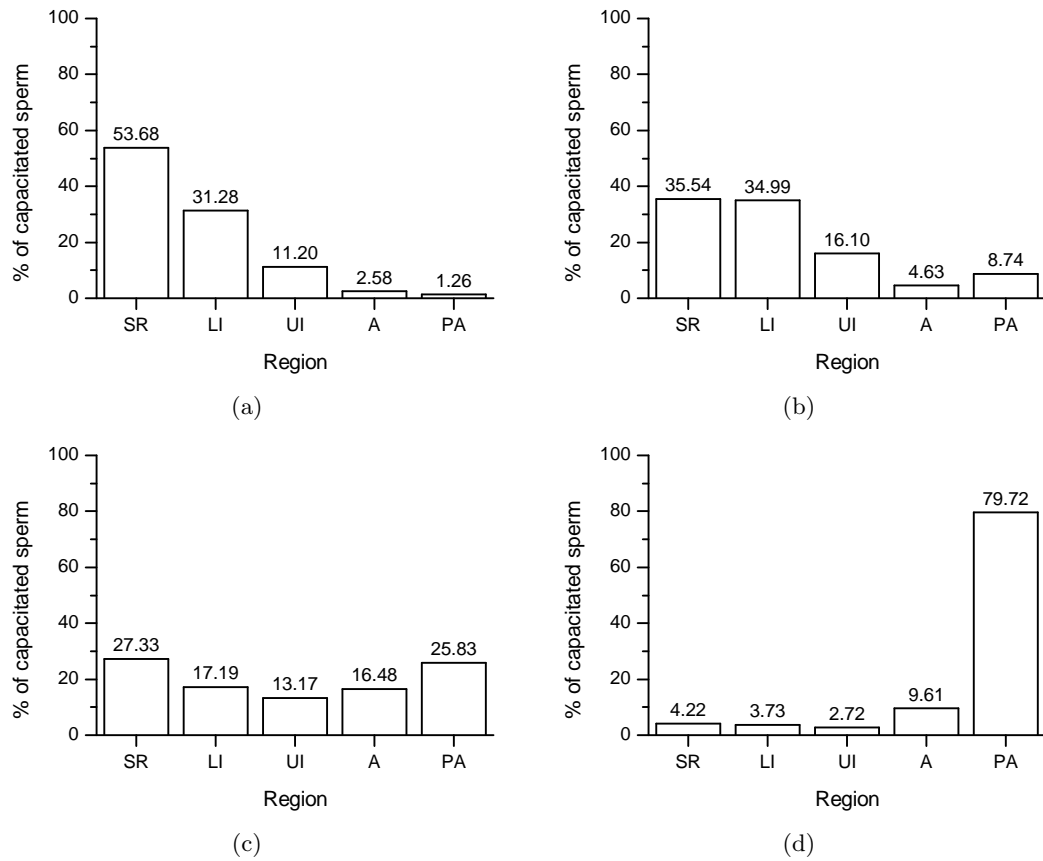
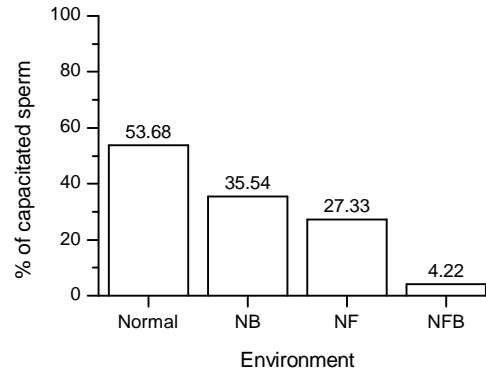


Figure 8.4: The distribution of sperm in (a) the normal environment, (b) the environment without bends, (c) then environment without folds and (d) the environment without bends or folds. SR = sperm reservoir, LI = lower isthmus, UI = upper isthmus, A = ampulla, PA = pre-ampulla.

### 8.1.2.1 Sperm reservoir

Figure 8.5 shows the distribution of sperm in the sperm reservoir at the end of the simulation as a percentage of the total capacitated population. Only sperm which have become capacitated during the simulation are included.



*Figure 8.5:* The relative percentage of capacitated sperm retained in the sperm reservoir under different environmental conditions. Normal = the normal environment, NB = the environment without bends, NF = the environment without folds, NFB = the environment without bends of folds.

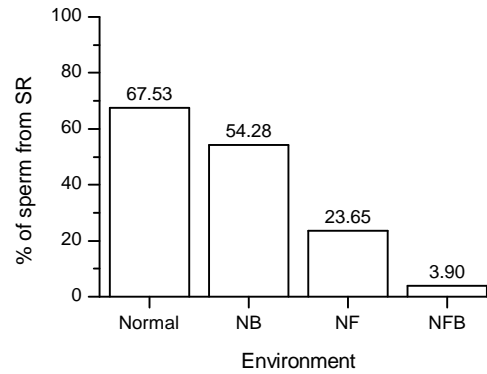
The results show that under normal conditions, around 54% of the capacitated sperm population are unable to leave the sperm reservoir. When the bends are removed, 18% fewer sperm are retained and when the folds are removed, 26% fewer sperm are retained. The combined removal of both bends and folds results in only 4% of sperm being retained in the sperm reservoir.

### 8.1.2.2 Lower isthmus

Figure 8.6 shows the distribution of sperm in the lower isthmus at the end of the simulation as a percentage of the number of sperm entering the lower isthmus from the sperm reservoir.

The results show that the removal of bends has a relatively low impact on the number of sperm retained, reducing it by 13% from the normal case. Removing the folds has a larger impact, and the combined removal of both bends and folds results in only 4% of sperm being retained in the lower isthmus.

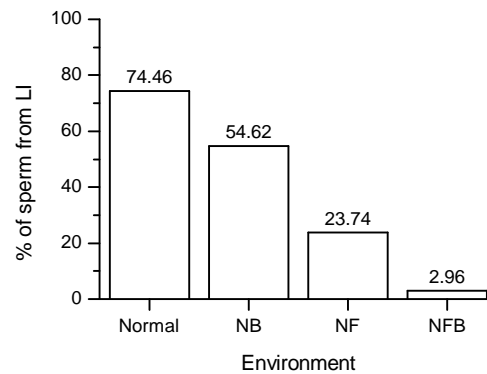




*Figure 8.6:* The relative percentage of sperm entering from the sperm reservoir which were retained in the lower isthmus under different environmental conditions. Normal = the normal environment, NB = the environment without bends, NF = the environment without folds, NFB = the environment without bends of folds.

### 8.1.2.3 Upper isthmus

Figure 8.7 shows the distribution of sperm in the upper isthmus at the end of the simulation as a percentage of the number of sperm entering the upper isthmus from the lower isthmus.

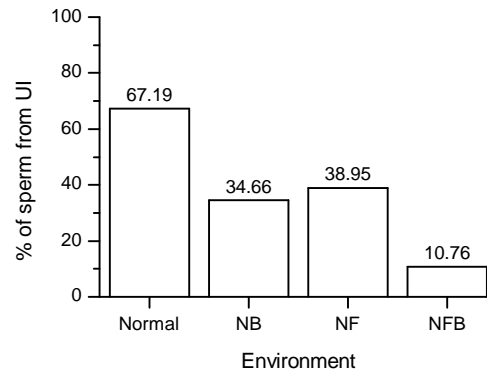


*Figure 8.7:* The relative percentage of sperm entering from the lower isthmus which were retained in the upper isthmus under different environmental conditions. Normal = the normal environment, NB = the environment without bends, NF = the environment without folds, NFB = the environment without bends of folds.

This shows a similar distribution to Figure 8.6. As both the lower isthmus and the upper isthmus has a similar 3D structure, this is to be expected.

#### 8.1.2.4 Ampulla

Figure 8.8 shows the distribution of sperm in the ampulla at the end of the simulation as a percentage of the number of sperm entering the ampulla from the upper isthmus.



*Figure 8.8:* The relative percentage of sperm entering from the upper isthmus which were retained in the ampulla under different environmental conditions. Normal = the normal environment, NB = the environment without bends, NF = the environment without folds, NFB = the environment without bends of folds.

When viewed as a percentage of sperm leaving the upper isthmus, the normal case is able to retain a larger percentage of sperm than the other cases. Unlike the other regions of the oviduct, in the ampulla the bends have a greater impact on the retention of sperm than the folds.

#### 8.1.2.5 Pre-ampulla

As all the sperm leaving the upper isthmus will either end up in the ampulla or the pre-ampulla, Figure 8.9 shows the distribution of sperm in the pre-ampulla at the end of the simulation as a percentage of the number of sperm leaving the upper isthmus.

This shows that for the normal environment, just over 30% of the sperm leaving the upper isthmus will pass through to the pre-ampulla. This figure is approximately doubled when the bends or folds are removed. When there are no bends or folds, the vast majority of sperm leaving the upper isthmus arrive in the pre-ampulla.

#### 8.1.3 Penetrated oocytes

Figure 8.10(a) shows the percentage of oocytes that were penetrated by the end of the simulation. For normal environments, 74% of oocytes were penetrated. For all modified environments, the average number of penetrated oocytes is almost 100%.

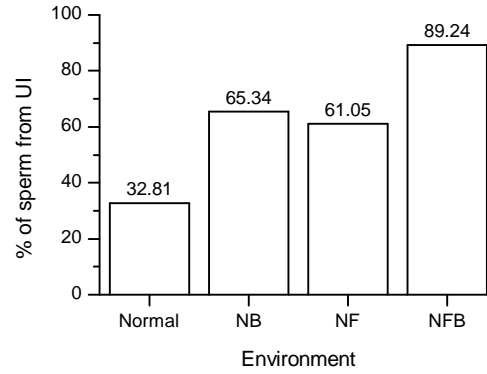


Figure 8.9: The relative percentage of sperm entering from the upper isthmus which passed through the ampulla and remained in the pre-ampulla under different environmental conditions. Normal = the normal environment, NB = the environment without bends, NF = the environment without folds, NFB = the environment without bends of folds.

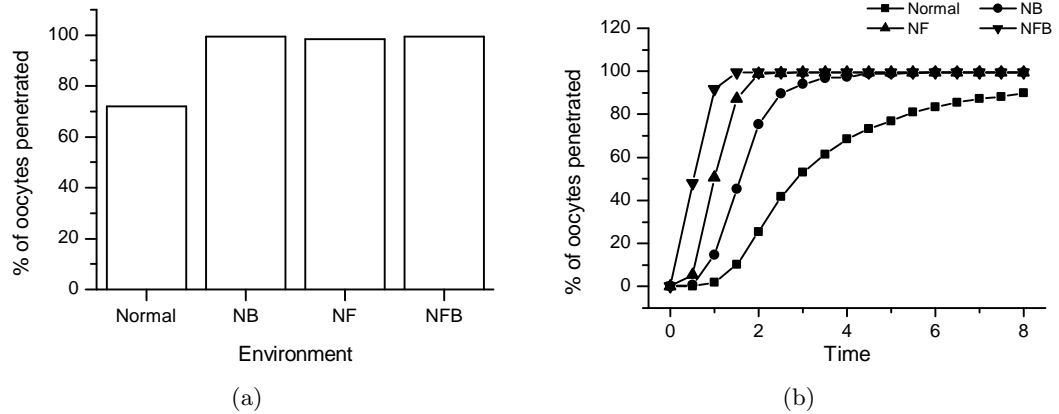
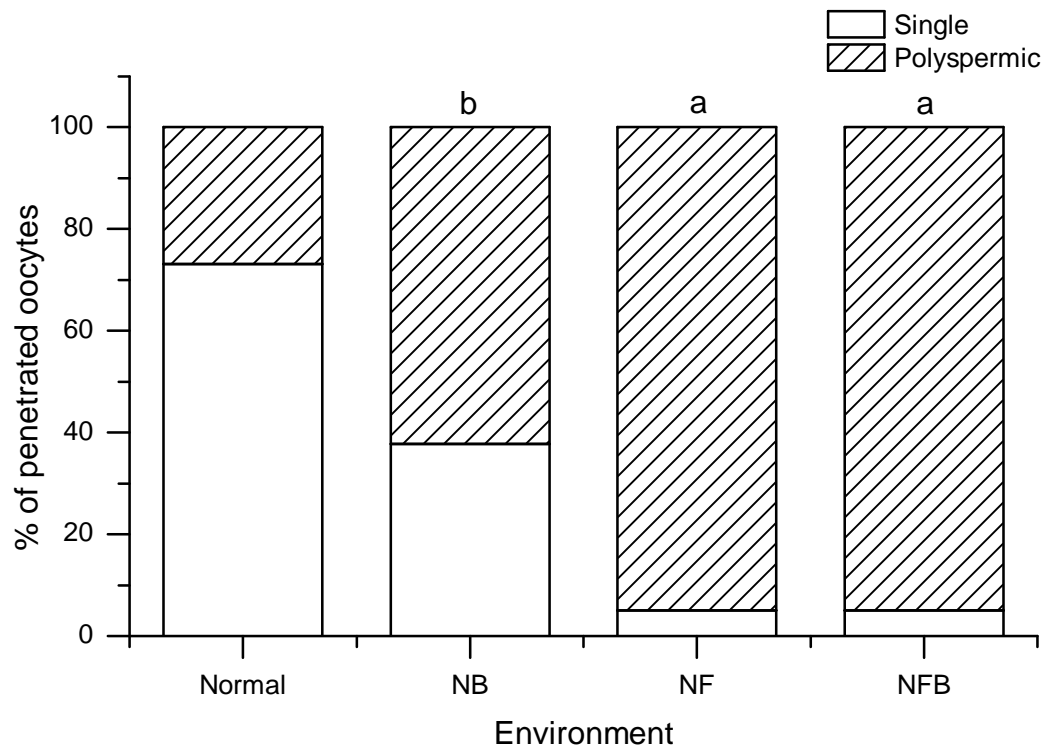


Figure 8.10: (a) The percentage of penetrated oocytes at the end of the simulation and (b) how the percentage of penetrated oocytes changes over time for the different environmental conditions. Normal = the normal environment, NB = the environment without bends, NF = the environment without folds, NFB = the environment without bends of folds.

This implies that removing the bends or curves has the same impact on the number of penetrated oocytes. Figure 8.10(b) shows the time period over which the oocytes are penetrated. This shows that with no folds or bends, then the maximum percentage of oocytes are penetrated in the shortest amount of time. Therefore, it appears that the presence of bends and folds actually hinders the reproductive process. However, when considering polyspermy, then this is no longer true.

### 8.1.4 Polyspermic oocytes

Figure 8.11 shows the relative percentage of single and polyspermic oocytes out of all penetrated oocytes.



*Figure 8.11:* The percentage of oocytes which are polyspermic out of all penetrated oocytes under different environmental conditions. Normal = the normal environment, NB = the environment without bends, NF = the environment without folds, NFB = the environment without bends of folds.

Although in the normal case only 72% of oocytes are penetrated, whereas in the modified environments closer to 97% are penetrated (Figure 8.10), a much higher percentage of the penetrated oocytes are polyspermic. It seems that a lack of bends increases polyspermy from 24% to 60% ,whereas a lack of folds increases it to around 95%.

The results presented in this section show that the structure of the oviduct and the combination of both the external bends and the internal folds is essential for the regulated passage of sperm to the site of fertilisation.

### 8.1.5 Experiment 1: Summary and conclusions

With regard to hypothesis (A) defined in Section 8.1, the removal of the bends or the folds resulted in a significant change in the percentage of penetrated oocytes. Therefore, the hypothesis is rejected. For hypothesis (B) the removal of the bends or the folds resulted a significant change in polyspermic oocytes, so this hypothesis is also rejected.

With regard to hypothesis (C), the bends play a less significant role than the folds, and removal of the bends still results in a reasonable number of oocytes with single sperm penetration. For the normal case, 52% of the total oocyte population are penetrated by just one sperm, whereas without bends 39% of the total oocyte population are penetrated by one sperm. Without folds, only 4% of the total oocyte population are penetrated by just one sperm. Therefore, the hypothesis is accepted.

The sexually transmitted disease Chlamydia, in its acute form, has been observed to cause the internal folds and the ciliated lining of the oviduct to disappear (Zana et al., 1991). It therefore likely that this lack of folds contributes significantly to the loss of fertility caused by the disease.

Therefore to answer the questions “does the complex 3D structure of the oviduct regulate the passage of sperm to the site of fertilisation?” and “Which aspects of the 3D model are most significant with respect to the regulation of sperm progression to the site of fertilisation?”, we can say that given the assumptions (Appendix D), the structure of the 3D environment does regulate the passage of sperm to the site of fertilisation in mice, with the internal folds playing a more significant role than the external bends. These findings are subject to the simplifications of the model when compared to the real system.

## 8.2 Experiment 2: sperm population size

Sperm counts in the oviduct vary significantly between individuals, breeds, strains, genera and species. However, successful fertilisation occurs in all. Correctly estimating the volume and therefore number of sperm required for optimal fertilisation is an important challenge for the livestock industry. It is also a concern for humans, as it has been shown that sperm counts and density have dropped significantly over the past few decades, which could lead to greater infertility in the population (Swan et al., 1997). In livestock and other mammals, studies have been carried out to find the optimal amount of sperm to inseminate into the uterus, which is often several million. However, there is no clear idea of how many sperm are actually required once they reach the oviduct.

This section attempts to address the research question posed in Chapter 1, (e) “How does the number of sperm within the oviduct affect the number of penetrated oocytes”.

To address research question (e), an exploratory study was performed to investigate the relationships between the initial sperm population size and sperm - oocyte interactions over time. An experiment was setup to investigate how the distribution of sperm, oocyte penetration and incidence of polyspermy change when the total sperm population is varied. The different sperm population sizes used as inputs into the system are shown in Table 8.1, relative to the average number of sperm used in all previous experiments, which is used as a control. A set of experiments was run for each sperm population size using all 9 environments with 40 replicates per simulation. The simulations were run for 12 hours instead of the normal 8. This will help to determine if a reduction in the number of sperm can result in a similar fertility rate over a longer period of time.

<b>No. of sperm</b>	<b>Relative to control</b>
3072	x2
2304	x1.5
1536	x1
768	x0.5
308	x0.2
154	x0.1
77	x0.05

*Table 8.1:* The different sperm population sizes used for the experiment. The factor differences from the average population size identified from literature is also shown.

Figure 8.12 shows the relative distribution of capacitated sperm within the oviduct at the end of the simulation. Regardless of the number of sperm in the system, the relative distribution remains the same. As the capacitation rate is based on a probability, these results are not surprising.

Figure 8.13 shows how the percentage of penetrated oocytes changes with different numbers of sperm. Penetration only starts after four hours *post coitus*. This shows that increasing the number of sperm by 50% results in a 10% increase in the number of penetrated oocytes at 8 hours, whereas reducing the sperm population by 50% results in a 20% decrease. From these results, it appears that doubling sperm population is beneficial as a larger number of oocytes are penetrated. However, if you consider polyspermy, then things are a little different.

Figure 8.14 shows the total number of oocytes which have been penetrated by a single sperm only, as a percentage of the total number of oocytes. From these results,

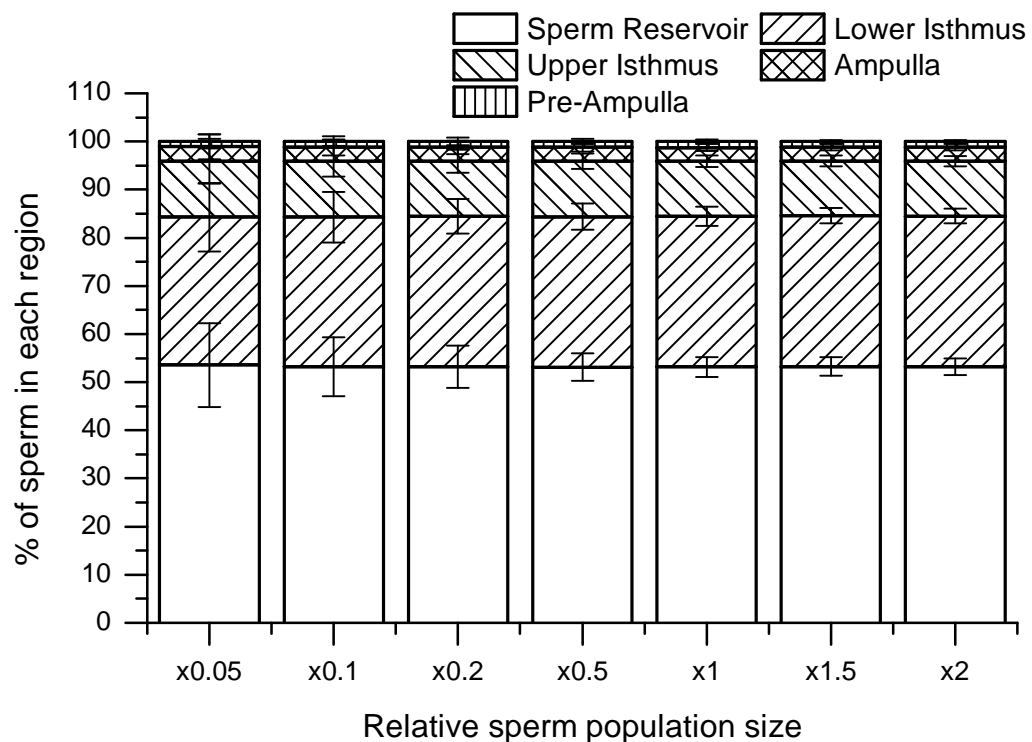


Figure 8.12: The relative distribution of sperm within the oviduct after 12 hours for the different sperm population sizes. Error bars show SD.

it appears that the control value of 1,536 sperm is ideal. Increasing the population by 50% and reducing it by 50% results in a similar number of oocytes with single sperm penetration. Doubling the number of sperm results in a significant reduction of monospermic oocytes.

Figure 8.15 shows how the number of polyspermic oocytes as a percentage of all penetrated oocytes changes over time. The relative number of polyspermic oocytes appears to stabilise after 7 or 8 hours. Increasing the sperm population by 50% results in approximately 50% greater incidence of polyspermy.

The results presented in this section show that even with large variations in the sperm population, successful fertilisation to some degree occurs.

There is a non-linear relationship between the number of sperm and the number of penetrated oocytes over time. Increasing the number of sperm by 50% results in a 10% increase in the number of penetrated oocytes at 8 hours and reducing the sperm population by 50% results in a 20% decrease in the number of penetrated oocytes at 8 hours. With respect to polyspermy, increasing the number of sperm by 50% results

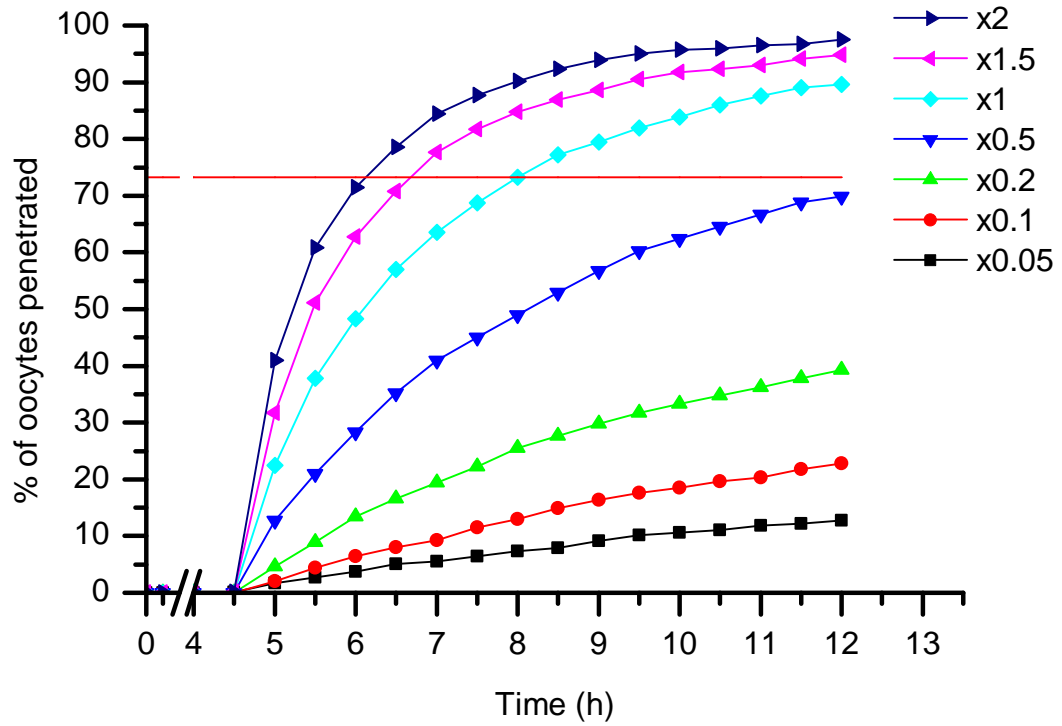


Figure 8.13: How the percentage of penetrated oocytes changes over time for the different sperm population sizes. The horizontal red line marks the average percentage of penetrated oocytes after 8 hours for the control sperm population.

in a significantly greater incidence of polyspermy. Decreasing the number of sperm by 50% results in significantly lower incidence of polyspermy. However, when the number of monospermic oocytes considered, increasing or decreasing the number of sperm by 50% results in a similar number of monospermic oocytes.

To answer the question “How does the number of sperm within the oviduct affect the number of penetrated oocytes?”, we can say that there is a non-linear relationship between the number of sperm in the oviduct and the number of penetrated oocytes, an increase in the number of sperm results in an increase in polyspermy and a decrease in the number of sperm results in a decrease in polyspermy. We can also say that large variations ( $\pm 50\%$ ) in the number of sperm in the oviduct can still result in a similar number of monospermic oocytes. This implies a degree of robustness to the system.

The results do imply that the average sperm population of 1536, which was the average for mice reported in literature (Hodges et al., 2008), is the ideal population size. However, these results could be an artefact of the model calibration process. As the model was calibrated using the sperm population in the ampulla after 8 hours, the



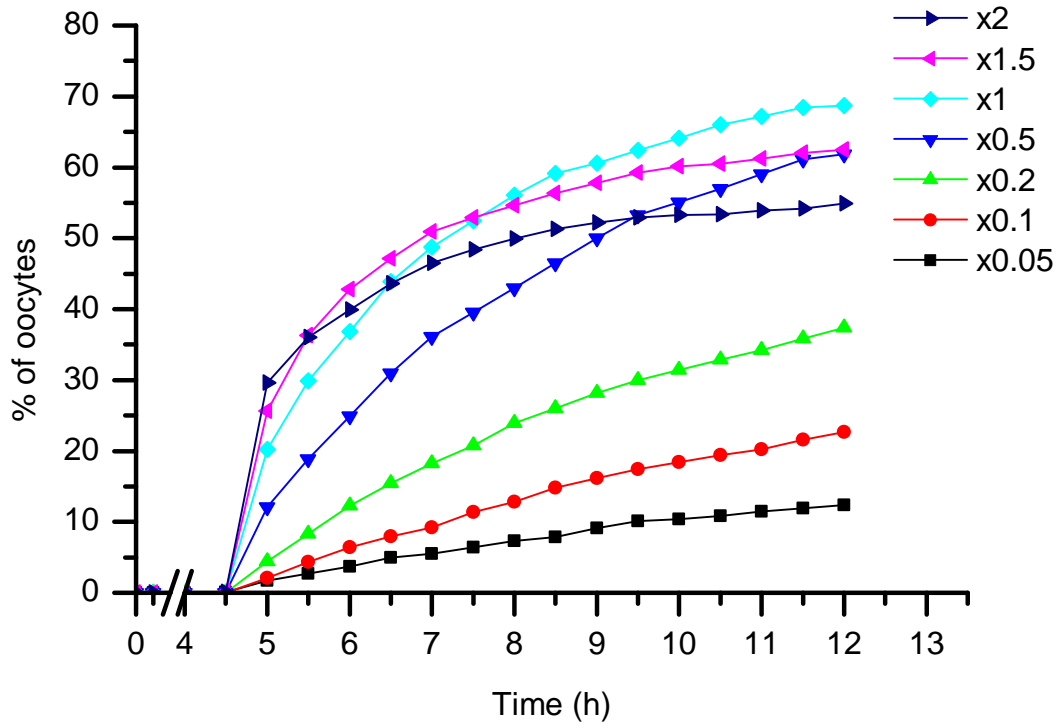


Figure 8.14: How the percentage of monospermic oocytes out of all oocytes changes over time for the different sperm population sizes.

ideal sperm population will be related to the number of sperm which have become capacitated within the 8 hour period. This is linked to the [TRY TO CAPACITATE] process, which is encoded as a percentage of the population over time. It is possible that the number of sperm which become capacitated over time could be more regulated, with a similar number of sperm becoming capacitated over time regardless of the total population size.

In literature, relationships between sperm population size, oocyte penetration and polyspermy *in vitro* have been identified (Fraser and Maudlin, 1978; Siddiquey and Cohen, 1982; Tsunoda and Chang, 1975). However, in the *in vitro* experiments, all sperm start in close proximity of the oocytes and the complexities of the oviductal environment are not represented.

Within the model, incidence of polyspermy appears to be mathematically related to the sperm population size and therefore the number of capacitated sperm over time. This is likely due to the fact that the number of capacitated sperm over time will determine the rate at which sperm reach the site of fertilisation. If the rate is increased, then

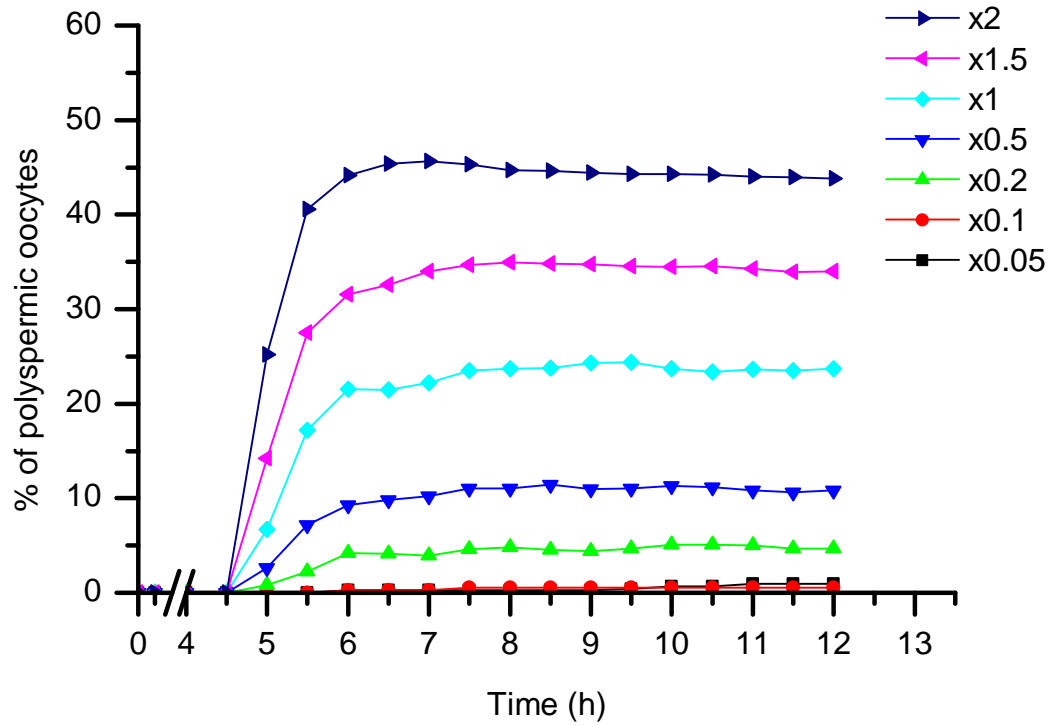


Figure 8.15: How the percentage of polyspermic oocytes out of all penetrated oocytes changes over time for the different sperm population sizes.

there is a greater chance of polyspermy.

### 8.3 Conclusions

The purpose of the chapter was to use the validated model of sperm behaviour to investigate the research questions posed in Chapter 1.

First, the impact that variations in the 3D environment have on the distribution of sperm, number of penetrated oocytes and incidence of polyspermy was investigated. The changes between individual environments showed a significant difference on sperm population distribution but not on oocyte penetration or polyspermy. When the impact of the 3D bends and internal folds were isolated, differences in the 3D bends were shown to have a significant impact on the sperm population distribution, but not differences in the 3D folds. Physically removing the bends or folds both had a significant impact on the distribution of sperm, oocyte penetration and polyspermy, with folds playing a more significant role than bends in the regulation of sperm passage to the site of fertilisation.

Second, the impact that an increase or reduction in the number of sperm in the oviduct on the distribution of sperm, penetrated oocytes and polyspermy was investigated. Regardless of the number of sperm in the oviduct, the relative distribution of sperm remained the same. After 8 hours, the percentage of polyspermic oocytes became constant, with a proportional difference relative to the population size. The results show that large variations in the number of sperm in the oviduct still result in successful fertilisation, and that a 50% increase and a 50% decrease in the sperm population can result in a similar number of monospermic oocytes.

## Chapter 9

# Conclusion

The purpose of this thesis was to investigate sperm movement in the female reproductive system and how the behaviour is influenced and regulated by the complex 3D structure of the oviduct.

The main research objectives were:

- 1 Review the biological literature surrounding sperm behaviour within the mammalian oviduct and identify the significant processes and parameters.
- 2 Review the computer science literature surrounding computational modelling of biological systems and identify the most suitable techniques for developing a computational model of sperm behaviour.
- 3 Develop a generic conceptual model of sperm navigation within the oviduct.
- 4 Construct a set of 3D computational models of the oviduct for use within the model.
- 5 Identify an agent based framework suitable for simulating large numbers of agents over a long time period.
- 6 Implement a computational model of sperm behaviour and identify and manage the data.
- 7 Develop a method for presenting the results of the model to biologists.
- 8 Validate the computational model of sperm behaviour against multiple literature sources.
- 9 Use the validated computational model to investigate the research questions.

Objective (1) was addressed in Chapters 2 and 4. In Chapter 2, a detailed description of mammalian reproduction was presented, the components of the reproductive system were identified and the different theories regarding sperm transport and navigation were discussed. In Chapter 4, the processes and parameters included in the model were specified and the spatial and temporal scales were defined. How these processes and parameters relate to each other and the identification of values for each parameter from literature is also described.

Objective (2) was addressed in Chapter 3. A review of the literature relating to computational modelling was presented, along with the commonly used techniques for modelling biological systems. Agent based modelling was identified as a suitable technique for modelling cellular behaviour. Existing computational models relating to the reproductive system were also presented.

Objective (3) was addressed in Chapter 4. The first agent based model of sperm behaviour in the mammalian oviduct was presented. The important aspects and behaviour of the model were identified and linked together. The mouse was identified as a target species for model validation and suitable values for all parameters were identified.

Objective (4) was addressed in Chapter 5. Existing techniques for constructing 3D environments from biological samples were found to be insufficient for capturing the small scale soft tissue folds of the oviduct. Therefore, a set of techniques were developed for identifying the curved bends of the mouse oviduct and the topological order of cross sections taken from histology images. A set of 3D models was then generated which contained representations of the external bends and internal folds, which are the two most significant structural features of the environment. The identified 3D curves were validated by visual comparison with the stacked source images in 3D. The internal folds were validated by visual comparison with the corresponding histology images for the different regions of the oviduct.

Objective (5) was addressed in Chapter 6. A selection of commonly used agent modelling frameworks were compared with FLAME and FLAME GPU, which were developed within the University of Sheffield. Initially, the frameworks were compared using StupidModel, using the descriptions presented by Railsback et al. (2006). A more comprehensive comparison using evaluation criteria specific to the model requirements was then performed. FLAME GPU was found to be superior in terms of both coding simplicity and performance with large and small numbers of agents.

Objective (6) was addressed in Chapter 6. The computational implementation of the conceptual model presented in Chapter 4 is described, along with technical challenges

such as optimal collision detection with the 3D environments constructed in Chapter 5. The model was modified within the framework to allow concurrent simulation replicates to be performed for statistically significant data. The data generated by the model was then aggregated to produce a meaningful set of results.

Objective (7) was addressed in Chapter 6. A visualisation tool was developed to allow the data to be presented visually. The 3D environment was displayed using a ‘dual depth peeling’ technique developed by Bavoil and Myers (2008), which allows 3D objects to be displayed with varying degrees of transparency. The position of the agents is displayed along with a line connecting the previous 5 positions, therefore showing the path that the agent has followed through the model.

Objective (8) was addressed in Chapter 7. The model was calibrated against literature sources which present the number of in the site of fertilisation over time. Only the parameters which were estimated or had uncertain values from literature were modified. Once a reasonable distribution of sperm over time was established, sensitivity analysis of the model was performed to investigate the impact of all model parameters in relation to the distribution of sperm and the number of oocyte penetrations. The 3D model was then modified to reduce the lumen swimming space by 10% and 20% to identify the impact that histology shrinkage may have on the model, and was found to impact the distribution of sperm within the oviduct, but not the number of penetrated oocytes. Finally, the model is validated against two independent studies reporting oocyte penetrations over time, and one study reporting polyspermy under normal conditions.

Objective (9) was addressed in Chapters 7 and 8. In Chapter 7, the relative significance of the parameters was investigated, and the progressive attachment threshold, the capacitation threshold and the life of sperm once capacitated all have a significant impact on the results. The basic mechanisms implemented were sufficient to accurately predict the number of penetrated oocytes over time. However, as many of the parameters were calibrated or estimated, it is possible that they are compensating for one or more of the more complex navigation mechanisms such as chemotaxis.

In Chapter 8, the complex structure of the oviduct was shown to play a key role in regulating the passage of sperm to the site of fertilisation. Natural variations in the 3D bends were shown to have a significant influence on the distribution of sperm, but not on the number of penetrated oocytes. The absence of folds resulted in very high incidence of polyspermy, whereas the absence of bends resulted double the percentage of polyspermic oocytes.

Increasing the number of sperm in the system resulted in an increase of oocyte penetrations and reducing the number of sperm resulted in a decrease. However, increasing the number of sperm resulted in a significant increase in polyspermy, with fewer single penetrated oocytes. With a reduced number of sperm, the oocyte penetration at 12 hours was similar to the normal number of sperm at 8 hours, although with significantly fewer polyspermic oocytes.

### 9.1 Limitations and future work

Computational modelling of the reproductive system is a relatively new field and there are many different areas for future investigation. In relation to the model of sperm behaviour, there are many different aspects that could be investigated.

The difficulties in obtaining accurate measurements for the parameters of the model were discussed in Chapter 4. Due to the differences between different species, breeds, strains and individuals, the values reported in literature have large variations. These differences are compounded when the different experimental conditions and measurement techniques are employed. In order to improve the model, the parameters which are estimated or calibrated could be made more realistic, specifically the angles that sperm turn when interacting with the oviduct and moving non-progressively. This could potentially be done by analysing video footage of sperm moving *in situ*, in a similar way to Demott and Suarez (1992). By measuring the angles between the direction sperm are swimming when they attach to the oviduct and the direction they swim away, a more accurate range of potential angles could be established. This type of measurement has to be performed *in situ*, as video footage *in vitro* would not show the physical constraints that the narrow lumen would have on the movement of sperm.

The model could be extended to provide a molecular level model of some of the processes, in particular attachment and detachment. By modelling the surface proteins that cause the sperm to attach to the oviduct and the loss of these proteins that occurs during capacitation, the resulting attachment and detachment behaviour could be made more realistic.

More complex mechanisms could also be introduced to see how they would influence the model. Finite element modelling could be used to establish a chemical or temperature gradient within the model, and the sperm agents could be modified to orient themselves towards the gradient during non-progressive movement. This would allow investigation of the conditions under which chemotaxis or thermotaxis would be feasible. Including fluid dynamics and the influence of peristaltic contractions on that fluid and the sperm within is currently not possible due to the massive computational power required to

fully resolve the fluid simulation within such a complex environment.

Local maternal communication could also be included as a factor. The walls of the oviduct could be modelled as small tiles, and collisions between sperm and the tiles could cause a change in the genetic expression at the oviductal location. This in turn could cause a change in the behaviour of subsequent sperm interacting with the same oviductal location.

The model could be translated to different species. Although a computational model of the fertility of mice does not have any obvious benefits to society, if the model was translated into other mammals, such as livestock or humans then the benefits become greater.

Heterogeneity in the sperm population could be increased. Individual sperm could have different movement speeds and individual parameters taken from CASA measurements. This would allow for the comparison between different sperm populations. For example, sperm from a very fertile male with high sperm motility could be compared with sperm from a male with low fertility and low sperm motility. This would potentially allow for the prediction of a successful oocyte penetration.

For humans, this would allow the tool to be used to aid assisted fertility treatments. It could calculate the probability of a successful pregnancy given CASA measurements and other factors from couples having trouble conceiving.



## Bibliography

- Alves, S. G., Martins, M. L., Fernandes, P. A. and Pittella, J. (1996), ‘Fractal patterns for dendrites and axon terminals’, *Physica A: Statistical and Theoretical Physics* **232**(1-2), 51–60.
- Amrani, M. and Shariat, B. (2000), Deformable organs modeling with multi layer particle systems, in ‘IEEE International Conference on Information Visualization’, IEEE Computer Society, Los Alamitos, CA, USA, pp. 351–356.
- Anderson, A. R. A., Chaplain, M. A. J., Newman, E. L., Steele, R. J. C. and Thompson, A. M. (2000), ‘Mathematical modelling of tumour invasion and metastasis’, *Journal of Theoretical Medicine* **2**(2), 129–154.
- Archip, N., Rohling, R., Dessenne, V., Erard, P.-J. and Nolte, L. P. (2006), ‘Anatomical structure modeling from medical images’, *Computer Methods and Programs in Biomedicine* **82**(3), 203–215.
- Athale, C., Mansury, Y. and Deisboeck, T. S. (2005), ‘Simulating the impact of a molecular ‘decision-process’ on cellular phenotype and multicellular patterns in brain tumors’, *Journal of Theoretical Biology* **233**(4), 469–481.
- Attur, M. G., Dave, M. N., Tsunoyama, K., Akamatsu, M., Kobori, M., Miki, J., Abramson, S. B., Katoh, M. and Amin, A. R. (2002), ‘“a system biology” approach to bioinformatics and functional genomics in complex human diseases: Arthritis’, *Current Issues in Molecular Biology* **4**, 129–146.
- Austin, C. R. (1952), ‘The ‘capacitation’ of the mammalian sperm’, *Nature* **170**(4321), 326.
- Austin, C. R. and Braden, A. W. H. (1956), ‘Early reactions of the rodent egg to spermatozoon penetration’, *Journal of Experimental Biology* **33**(2), 358–365.
- Auyang, S. Y. (1998), *Foundations of Complex-system Theories: In Economics, Evolutionary Biology, and Statistical Physics*, Cambridge University Press.

- Bahat, A., Tur-Kaspa, I., Gakamsky, A., Giojalas, L. C., Breitbart, H. and Eisenbach, M. (2003), ‘Thermotaxis of mammalian sperm cells: a potential navigation mechanism in the female genital tract.’, *Nature Medicine* **9**(2), 149–150.
- Bailey, A. M., Lawrence, M. B., Shang, H., Katz, A. J. and Peirce, S. M. (2009), ‘Agent-Based model of therapeutic Adipose-Derived stromal cell trafficking during ischemia predicts ability to roll on P-Selectin’, *PLoS Comput Biol* **5**(2), e1000294+.
- Bailey, A. M., Thorne, B. C. and Peirce, S. M. (2007), ‘Multi-cell agent-based simulation of the microvasculature to study the dynamics of circulating inflammatory cell trafficking’, *Annals of Biomedical Engineering* **35**(6), 916–936.
- Baker, S. J. (2006), ‘Steve’s portable game library’. [Last Accessed: 26-September-2011].  
**URL:** <http://plib.sourceforge.net>
- Bathe, K.-J. (2007), Finite element method, in B. W. Wah, ed., ‘Wiley Encyclopedia of Computer Science and Engineering’, John Wiley & Sons, Inc., Hoboken, NJ, USA.
- Battalia, D. E. and Yanagimachi, R. (1979), ‘Enhanced and co-ordinated movement of the hamster oviduct during the periovulatory period’, *Journal of Reproduction and Fertility* **56**(2), 515–520.
- Bavoil, L. and Myers, K. (2008), Order independent transparency with dual depth peeling, Technical report, NVIDIA Developer SDK 10.
- Beck, L. R. and Boots, L. R. (1972), The comparative anatomy, histology, and morphology of the mammalian oviduct, in A. D. Johnson and C. W. Foley, eds, ‘The oviduct and its functions’, Academic Press, New York, pp. 1–52.
- Bedford, J. M. (1983), ‘Significance of the need for sperm capacitation before fertilization in eutherian mammals.’, *Biology of Reproduction* **28**(1), 108–120.
- Benitz, A. M. and Graves, C. N. (1972), ‘In vitro aging of mouse spermatozoa’, *Journal of Animal Science* **35**(1), 236–237.
- Bennett, D. and Dunn, L. C. (1967), ‘Studies of effects of t-alleles in the house mouse on spermatozoa’, *Journal of Reproductive Fertility* **13**(3), 421–428.
- Bergman, D. L. and Ullberg, U. (1998), ‘Scaling properties of the placenta’s arterial tree’, *Journal of Theoretical Biology* **193**(4), 731–738.

- Bobashev, G. V., Goedecke, M. D., Yu, F. and Epstein, J. M. (2007), A hybrid epidemic model: Combining the advantages of Agent-Based and Equation-Based approaches, in 'Simulation Conference, 2007 Winter', IEEE, pp. 1532–1537.
- Bonabeau, E. (2002), 'Agent-based modeling: Methods and techniques for simulating human systems', *Proceedings of the National Academy of Sciences of the United States of America* **99**(Suppl 3), 7280–7287.
- Bottino, D., Mogilner, A., Roberts, T., Stewart, M. and Oster, G. (2002), 'How nematode sperm crawl', *Journal of Cell Science* **115**(2), 367–384.
- Box, G. E. P. and Draper, N. R. (1986), *Empirical model-building and response surface*, John Wiley & Sons, Inc., New York, NY, USA.
- Braden, A. W., Austin, C. R. and David, H. A. (1954), 'The reaction of zona pellucida to sperm penetration.', *Australian journal of biological sciences* **7**(3), 391–409.
- Braden, A. W. H. (1958), 'Variation between strains of mice in phenomena associated with sperm penetration and fertilization', *Journal of Genetics* **56**(1), 37–47.
- Braden, A. W. H. (1959), 'Strain differences in the morphology of the gametes of the mouse', *Australian Journal of Biological Sciences* **12**(1), 65–71.
- Braden, A. W. H. and Austin, C. R. (1954a), 'Fertilization of the mouse egg and the effect of delayed coitus and of Hot-Shock treatment.', *Australian Journal of Biological Sciences* **7**(4), 552–565.
- Braden, A. W. H. and Austin, C. R. (1954b), 'The number of sperms about the eggs in mammals and its significance for normal fertilization', *Australian Journal of Biological Sciences* **7**(4), 543–551.
- Braden, A. W. H. and Gluecksohn-Waelsch, S. (1958), 'Further studies of the effect of the t locus in the house mouse on male fertility', *Journal of Experimental Zoology* **138**(3), 431–452.
- Breen, D. E. (2000), A survey of cloth modeling methods, in D. H. House and D. E. Breen, eds, 'Cloth modeling and animation', A. K. Peters, Ltd., Natick, MA, USA, pp. 19–53.
- Brock, K. K., Sharpe, M. B., Dawson, L. A., Kim, S. M. and Jaffray, D. A. (2005), 'Accuracy of finite element model-based multi-organ deformable image registration', *Medical Physics* **32**(6), 1647–1659.

- Brokaw, C. J. (2001), ‘Simulating the effects of fluid viscosity on the behaviour of sperm flagella’, *Mathematical Methods in the Applied Sciences* **24**(17-18), 1351–1365.
- Brokaw, C. J. and Luck, D. J. L. (1983), ‘Bending patterns of chlamydomonas flagella i. wild-type bending patterns’, *Cell Motility* **3**(2), 131–150.
- Burkitt, M. (2008), Procedural modelling of an oviduct, Master’s thesis, Department of Computer Science, University of Sheffield, Sheffield, UK.
- Burkitt, M., Romano, D. M., Walker, D. and Fazeli, A. (2010), 3D modelling of complex biological structures: The oviduct, in ‘EG UK Theory and Practice of Computer Graphics’, Eurographics Association, University of Sheffield, UK, pp. 255–262.
- Butcher, E. C., Berg, E. L. and Kunkel, E. J. (2004), ‘Systems biology in drug discovery’, *Nature Biotechnology* **22**(10), 1253–1259.
- Caiazzo, A., Evans, D., Falcone, J.-L., Hegewald, J., Lorenz, E., Stahl, B., Wang, D., Bernsdorf, J., Chopard, B., Gunn, J., Hose, R., Krafczyk, M., Lawford, P., Smallwood, R., Walker, D. and Hoekstra, A. (2011), ‘A complex automata approach for in-stent restenosis: Two-dimensional multiscale modelling and simulations’, *Journal of Computational Science* **2**(1), 9–17.
- Carley, K. (1996), Validating computational models, Technical report, Carnegie Mellon University.
- Chaichana, T., Sun, Z. and Jewkes, J. (2011), ‘Computation of hemodynamics in the left coronary artery with variable angulations’, *Journal of Biomechanics* **44**(10), 1869–1878.
- Champlin, A. K., Dorr, D. L. and Gates, A. H. (1973), ‘Determining the stage of the estrous cycle in the mouse by the appearance of the vagina’, *Biology of Reproduction* **8**, 491–494.
- Chilvers, M. A. and O’Callaghan, C. (2000), ‘Analysis of ciliary beat pattern and beat frequency using digital high speed imaging: comparison with the photomultiplier and photodiode methods’, *Thorax* **55**(4), 314–317.
- Clayton, R. H. and Panfilov, A. V. (2008), ‘A guide to modelling cardiac electrical activity in anatomically detailed ventricles’, *Progress in Biophysics and Molecular Biology* **96**(1-3), 19–43.
- Coakley, S. (2007), Formal Software Architecture for Agent-Based Modelling in Biology, PhD thesis, Department of Computer Science, University of Sheffield, UK.

- Coakley, S., Smallwood, R. and Holcombe, M. (2006), Using x-machines as a formal basis for describing agents in agent-based modelling, *in* 'Proceedings of 2006 Spring Simulation Multiconference', pp. 33–40.
- Cohen-Dayag, A., Tur-Kaspa, I., Dor, J., Mashiach, S. and Eisenbach, M. (1995), 'Sperm capacitation in humans is transient and correlates with chemotactic responsiveness to follicular factors', *Proceedings of the National Academy of Sciences of the United States of America* **92**(24), 11039–11043.
- Craciunescu, O., Das, S. K. and Dewhirst, M. K. (1999), 'Three-dimensional microvascular networks fractal structure: A potential for tissue characterization', *Advances in Heat and Mass Transfer in Biotechnology, HTD* **363**, 9–13.
- Crane, K., Llamas, I. and Tariq, S. (2007), Real-Time simulation and rendering of 3D fluids, *in* H. Nguyen, ed., 'GPU Gems 3', Addison Wesley Professional, chapter 30.
- Cummins, J. M. and Woodall, P. F. (1985), 'On mammalian sperm dimensions', *Journal of Reproductive Fertility* **75**(1), 153–175.
- Dagum, L. and Menon, R. (1998), 'OpenMP: an industry standard API for shared-memory programming', *Computational Science & Engineering, IEEE* **5**(1), 46–55.
- Dauplain, A., Favier, J. and Bottaro, A. (2008), 'Hydrodynamics of ciliary propulsion', *Journal of Fluids and Structures* **24**(8), 1156–1165.
- Debbaut, C., Monbaliu, D., Casteleyn, C., Cornillie, P., Van Loo, D., Masschaele, B., Pirenne, J., Simoens, P., Van Hoorebeke, L. and Segers, P. (2011), 'From vascular corrosion cast to electrical analog model for the study of human liver hemodynamics and perfusion', *IEEE Transactions on Biomedical Engineering* **58**(1), 25–35.
- Demott, R. P. and Suarez, S. S. (1992), 'Hyperactivated sperm progress in the mouse oviduct.', *Biology of Reproduction* **46**(5), 779–785.
- Dillon, R. and Fauci, L. J. (2000), 'An integrative model of internal axoneme mechanics and external fluid dynamics in ciliary beating', *Journal of Theoretical Biology* **207**(3), 415–430.
- Dillon, R. H., Fauci, L. J., Omoto, C. and Yang, X. (2007), 'Fluid dynamic models of flagellar and ciliary beating', *Annals of the New York Academy of Sciences* **1101**(Reproductive Biomechanics), 494–505.
- d'Inverno, M. and Saunders, R. (2005), Agent-Based modelling of stem cell self-organisation in a niche engineering Self-Organising systems, *in* S. A. Brueckner,

- G. Marzo Serugendo, A. Karageorgos and R. Nagpal, eds, 'Engineering SelfOrganising Systems Methodologies And Applications', Vol. 3464 of *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, Berlin, Heidelberg, chapter 4, pp. 52–68.
- Draghici, S., Khatri, P., Tarca, A. L., Amin, K., Done, A., Voichita, C., Georgescu, C. and Romero, R. (2007), 'A systems biology approach for pathway level analysis', *Genome Research* **17**(10), 1537–1545.
- Dukelow, W. R. and Riegler, G. D. (1972), Transport of gametes and survival of the ovum as functions of the oviduct, in A. D. Johnson and C. W. Foley, eds, 'The oviduct and its functions', Academic Press, New York, chapter 7, pp. 193–220.
- Eisenbach, M. and Giojalas, L. C. (2006), 'Sperm guidance in mammals an unpaved road to the egg', *Nature Reviews Molecular Cell Biology* **7**(4), 276–285.
- Epstein, J. M. (2009), 'Modelling to contain pandemics', *Nature* **460**(7256), 687.
- Farmer, J. D. and Foley, D. (2009), 'The economy needs agent-based modelling', *Nature* **460**(7256), 685–686.
- Fauci, L. J. and Dillon, R. (2006), 'Biofluidmechanics of reproduction', *Annual Review of Fluid Mechanics* **38**(1), 371–394.
- Fazeli, A. and Pewsey, E. (2008), 'Maternal communication with gametes and embryos: a complex interactome', *Briefings in Functional Genomics and Proteomics* **7**(2), 111–118.
- Firman, R. C. and Simmons, L. W. (2010), 'Sperm midpiece length predicts sperm swimming velocity in house mice', *Biology Letters* **6**(4), 513–516.
- Flugge, A. J., Timmis, J., Andrews, P., Moore, J. and Kaye, P. (2009), Modelling and simulation of granuloma formation in visceral leishmaniasis, in '2009 IEEE Congress on Evolutionary Computation (CEC)', IEEE, pp. 3052–3059.
- Foo, J. Y. and Lim, C. S. (2008), 'Biofluid mechanics of the human reproductive process: modelling of the complex interaction and pathway to the oocytes', *Zygote* **16**(04), 343–354.
- Foote, R. (2007), 'Mathematics and complex systems', *Science* **318**(5849), 410–412.
- Fraser, L. R. and Maudlin, I. (1978), 'Relationship between sperm concentration and the incidence of polyspermy in mouse embryos fertilized in vitro', *Journal of Reproduction and Fertility* **52**(1), 103–106.

- Friedrich, B. M. and Jülicher, F. (2007), ‘Chemotaxis of sperm cells.’, *Proceedings of the National Academy of Sciences of the United States of America* **104**(33), 13256–13261.
- Friedrich, B. M., Riedel-Kruse, I. H., Howard, J. and Jülicher, F. (2010), ‘High-precision tracking of sperm swimming fine structure provides strong test of resistive force theory’, *Journal of Experimental Biology* **213**(8), 1226–1234.
- Gabrys, E., Rybaczuk, M. and Kedzia, A. (2005), ‘Fractal models of circulatory system. symmetrical and asymmetrical approach comparison’, *Chaos, Solitons & Fractals* **24**(3), 707–715.
- Garcia-Alonso, A., Serrano, N. and Flaquer, J. (1994), ‘Solving the collision detection problem’, *Computer Graphics and Applications, IEEE* **14**(3), 36–43.
- Georgiou, A. S., Snijders, A. P., Sostaric, E., Aflatoonian, R., Vazquez, J. L., Vazquez, J. M., Roca, J., Martinez, E. A., Wright, P. C. and Fazeli, A. (2007), ‘Modulation of the oviductal environment by gametes.’, *Journal of proteome research* **6**(12), 4656–4666.
- Gil, M., Garcia-Herreros, M., Baron, F., Aparicio, I., Santos, A. and Garciamarin, L. (2009), ‘Morphometry of porcine spermatozoa and its functional significance in relation with the motility parameters in fresh semen’, *Theriogenology* **71**(2), 254–263.
- Gilbert, N. and Terna, P. (2000), ‘How to build and use agent-based models in social science’, *Mind & Society* **1**(1), 57–72.
- Giojalas, L. C., Rovasio, R. A., Fabro, G., Gakamsky, A. and Eisenbach, M. (2004), ‘Timing of sperm capacitation appears to be programmed according to egg availability in the female genital tract\*1’, *Fertility and Sterility* **82**(1), 247–249.
- Goodson, S. G., Zhang, Z., Tsuruta, J. K., Wang, W. and O’Brien, D. A. (2011), ‘Classification of mouse sperm motility patterns using an automated multiclass support vector machines model’, *Biology of Reproduction* **84**(6), 1207–1215.
- Gray, J. (1929), ‘The mechanism of ciliary movement’, *The American Naturalist* **63**(684), 68–81.
- Gray, J. and Hancock, G. J. (1955), ‘The propulsion of Sea-Urchin spermatozoa’, *Journal of Experimental Biology* **32**(4), 802–814.
- Grimm, V. (1994), ‘Mathematical models and understanding in ecology’, *Ecological Modelling* **75–76**, 641–651.

- Grimm, V. (1999), 'Ten years of individual-based modelling in ecology: what have we learned and what could we learn in the future?', *Ecological Modelling* **115**(2-3), 129–148.
- Grimm, V., Frank, K., Jeltsch, F., Brandl, R., Uchmaski, J. and Wissel, C. (1996), 'Pattern-oriented modelling in population ecology', *Science of The Total Environment* **183**(1-2), 151–166.
- Grzmil, P., Golas, A., Muller, C. and Styrna, J. (2007), 'The influence of the deletion on the long arm of the y chromosome on sperm motility in mice', *Theriogenology* **67**(4), 760–766.
- Gueron, S. and Levit-Gurevich, K. (2001), 'A three-dimensional model for ciliary motion based on the internal  $9 + 2$  structure', *Proceedings of the Royal Society B: Biological Sciences* **268**(1467), 599–607.
- Gueron, S. and Liron, N. (1993), 'Simulations of three-dimensional ciliary beats and cilia interactions.', *Biophysical journal* **65**(1), 499–507.
- Hafez, E. S. E. and Black, D. L. (1969), The mammalian uterotubal junction, in E. S. E. Hafez and R. J. Blandau, eds, 'The Mammalian Oviduct', The University of Chicago Press, Chicago, chapter 4, pp. 85–126.
- Harada, T. (2007), Real-Time rigid body simulation on GPUs, in H. Nguyen, ed., 'GPU Gems 3', Addison Wesley Professional, chapter 29.
- Harper, M. J. K. (1973), 'Relationship between sperm transport and penetration of eggs in the rabbit oviduct', *Biology of Reproduction* **8**(4), 441–450.
- Harper, M. J. K. (1982), Sperm and egg transport, in C. R. Austin and R. V. Short, eds, 'Germ cells and fertilization', second edn, University Press, Cambridge, chapter 5, pp. 102–127.
- Henein, C. M. and White, T. (2005), Agent-Based modelling of forces in crowds Multi-Agent and Multi-Agent-based simulation, in P. Davidsson, B. Logan and K. Takadama, eds, 'Multi-Agent and Multi-Agent-Based Simulation', Vol. 3415 of *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, Berlin, Heidelberg, chapter 14, pp. 173–184.
- Ho, H. C. and Suarez, S. S. (2001), 'Hyperactivation of mammalian spermatozoa: function and regulation', *Reproduction* **122**(4), 519–526.



- Ho, K., Wolff, C. A. and Suarez, S. S. (2009), 'CatSper-null mutant spermatozoa are unable to ascend beyond the oviductal reservoir.', *Reproduction, Fertility, and Development* **21**(2), 345–350.
- Hodges, C. A., Palmert, M. R. and Drumm, M. L. (2008), 'Infertility in females with cystic fibrosis is multifactorial: Evidence from mouse models', *Endocrinology* **149**(6), 2790–2797.
- Hogan, B., Beddington, R., Constantini, F. and Lacy, E. (1994), *Manipulating the Mouse Embryo*, 2nd edn, Cold Spring Harbor Laboratory Press, NY.
- Holdsworth, D. W. and Thornton, M. M. (2002), 'Micro-CT in small animal and specimen imaging', *Trends in Biotechnology* **20**(8), S34–S39.
- Holt, W. V. and Fazeli, A. (2010), 'The oviduct as a complex mediator of mammalian sperm function and selection', *Molecular Reproduction and Development* **77**(11), 934–943.
- Hood, L., Heath, J. R., Phelps, M. E. and Lin, B. (2004), 'Systems biology and new technologies enable predictive and preventative medicine.', *Science (New York, N.Y.)* **306**(5696), 640–643.
- Hornberg, J. J., Bruggeman, F. J., Westerhoff, H. V. and Lankelma, J. (2006), 'Cancer: A systems biology disease', *Biosystems* **83**(2-3), 81–90.
- Humphries, S., Evans, J. P. and Simmons, L. W. (2008), 'Sperm competition: linking form to function', *BMC Evolutionary Biology* **8**(1), 319+.
- Hunter, R. H. F. (1981), 'Sperm transport and reservoirs in the pig oviduct in relation to the time of ovulation.', *Journal of Reproduction and Fertility* **63**(1), 109–117.
- Hunter, R. H. F. (2008), 'Sperm release from oviduct epithelial binding is controlled hormonally by peri-ovulatory graafian follicles', *Molecular Reproduction and Development* **75**(1), 167–174.
- Hunter, R. H. F. and Léglise, P. C. (1971), 'Tubal surgery in the rabbit: Fertilization and polyspermy after resection of the isthmus', *American Journal of Anatomy* **132**(1), 45–52.
- Hyakutake, T., Hashimoto, Y., Yanase, S., Matsuura, K. and Naruse, K. (2009), 'Application of a numerical simulation to improve the separation efficiency of a sperm sorter', *Biomedical Microdevices* **11**(1), 25–33.

- Ideker, T., Galitski, T. and Hood, L. (2001), ‘A NEW APPROACH TO DECODING LIFE: Systems biology’, *Annual Review of Genomics and Human Genetics* **2**(1), 343–372.
- Ikawa, M., Inoue, N., Benham, A. M. and Okabe, M. (2010), ‘Fertilization: a sperm’s journey to and interaction with the oocyte’, *Journal of Clinical Investigation* **120**(4), 984–994.
- Ishikawa, M., Tsutsui, H., Cosson, J., Oka, Y. and Morisawa, M. (2004), ‘Strategies for sperm chemotaxis in the siphonophores and ascidians: A numerical simulation study’, *The biological bulletin* **206**(2), 95–102.
- Jackson, D. E., Holcombe, M. and Ratnieks, F. L. (2004), ‘Trail geometry gives polarity to ant foraging networks’, *Nature* **432**(7019), 907–909.
- Jacob, C., Litorco, J. and Lee, L. (2004), ‘Immunity through swarms: Agent-Based simulations of the human immune system artificial immune systems’, *Artificial Immune Systems* **3239**, 400–412.
- Jaiswal, B. S. and Eisenbach, M. (2002), Capacitation, in D. M. Hardy, ed., ‘Fertilization’, Academic Press, San Diego, California, USA, chapter 1, pp. 3–28.
- Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N. and Barabási, A. L. (2000), ‘The large-scale organization of metabolic networks.’, *Nature* **407**(6804), 651–654.
- Jin, M., Fujiwara, E., Kakiuchi, Y., Okabe, M., Satouh, Y., Baba, S. A., Chiba, K. and Hirohashi, N. (2011), ‘Most fertilizing mouse spermatozoa begin their acrosome reaction before contact with the zona pellucida during in vitro fertilization’, *Proceedings of the National Academy of Sciences* **108**(12), 4892–4896.
- Johnson, M. H. and Everitt, B. (1999), *Essential Reproduction (Essentials)*, 5th edition edn, WileyBlackwell.
- Katagiri, F. (2003), ‘Attacking complex problems with the power of systems biology’, *Plant Physiology* **132**(2), 417–419.
- Khronos, T. (2009), ‘Opencl - the open standard for parallel programming of heterogeneous systems’. [Last Accessed: 26-September-2011].  
**URL:** <http://www.khronos.org/opencl/>
- Kim, D.-H., Ko, D.-S., Lee, H.-C., Lee, H.-J., Park, W.-I., Kim, S. S., Park, J.-K., Yang, B.-C., Park, S.-B., Chang, W.-K. and Lee, H.-T. (2004), ‘Comparison of maturation, fertilization, development, and gene expression of mouse oocytes grown in vitro and in vivo’, *Journal of Assisted Reproduction and Genetics* **21**(7), 233–240.

- Kirk, D. B. and Hwu, W.-m. W. (2010), *Programming Massively Parallel Processors: A Hands-on Approach (Applications of GPU Computing Series)*, 1 edn, Morgan Kaufmann.
- Kitano, H. (2002), 'Systems biology: A brief overview', *Science* **295**(5560), 1662–1664.
- Kitaoka, H., Takaki, R. and Suki, B. (1999), 'A three-dimensional model of the human airway tree', *Journal of Applied Physiology* **87**(6), 2207–2217.
- Kleijnen, J. P. C. (1999), Validation of models: statistical techniques and data availability, in 'Simulation Conference Proceedings, 1999 Winter', Vol. 1, IEEE, pp. 647–654 vol.1.
- Kleijnen, J. (1995), 'Verification and validation of simulation models', *European Journal of Operational Research* **82**(1), 145–162.
- Klügl, F. (2008), A validation methodology for agent-based simulations, in 'Proceedings of the 2008 ACM symposium on Applied computing', SAC '08, ACM, New York, NY, USA, pp. 39–43.
- Kruger, J., Kipfer, P., Kondratieva, P. and Westermann, R. (2005), 'A particle system for interactive visualization of 3D flows', *IEEE Transactions on Visualization and Computer Graphics* **11**(6), 744–756.
- Law, A. M. (2005), How to build valid and credible simulation models, in 'WSC '05: Proceedings of the 37th conference on Winter simulation', Winter Simulation Conference, pp. 24–32.
- Lefebvre, R. and Suarez, S. S. (1996), 'Effect of capacitation on bull sperm binding to homologous oviductal epithelium.', *Biology of Reproduction* **54**(3), 575–582.
- Li, G., Citrin, D., Miller, R. W., Camphausen, K., Mueller, B., Mychalczak, B. and Song, Y. (2008), 3D and 4D medical image registration combined with image segmentation and visualization, in N. Wickramasinghe and E. Geisler, eds, 'Encyclopedia of Healthcare Information Systems', IGI Global, Hershey, PA, pp. 1–9.
- Luke, S., Cioffi-revilla, C., Panait, L. and Sullivan, K. (2004), Mason: A new multi-agent simulation toolkit, in 'Proceedings of the 2004 SwarmFest Workshop', Vol. 8.
- Lyons, R. A., Saridogan, E. and Djahanbakhch, O. (2006), 'The reproductive significance of human fallopian tube cilia', *Human Reproduction Update* **12**(4), 363–372.
- Malo, A. F., Gomendio, M., Garde, J., Lang-Lenton, B., Soler, A. J. and Roldan, E. R. S. (2006), 'Sperm design and sperm function', *Biology Letters* **2**(2), 246–249.

- Maluchnik, D. and Borsuk, E. (1994), ‘Sperm entry into fertilised mouse eggs’, *Zygote* **2**(02), 129–131.
- Mansury, Y., Kimura, M., Lobo, J. and Deisboeck, T. S. (2002), ‘Emerging patterns in tumor systems: Simulating the dynamics of multicellular clusters with an agent-based spatial agglomeration model’, *Journal of Theoretical Biology* **219**(3), 343–370.
- Marston, J. H. and Chang, M. C. (1964), ‘The fertilizable life of ova and their morphology following delayed insemination in mature and immature mice’, *J. Exp. Zool.* **155**(2), 237–251.
- Marzo, A., Singh, P., Larrabide, I., Radaelli, A., Coley, S., Gwilliam, M., Wilkinson, I. D., Lawford, P., Reymond, P., Patel, U., Frangi, A. and Hose, R. R. (2011), ‘Computational hemodynamics in cerebral aneurysms: The effects of modeled versus measured boundary conditions’, *Annals of Biomedical Engineering* **39**(2), 884–896.
- Mastroianni Jr, L. (1999), ‘The fallopian tube and reproductive health’, *Journal of Pediatric and Adolescent Gynecology* **12**(3), 121–126.
- Mazzilli, F., Rossi, T., Delfino, M., Dondero, F. and Makler, A. (2001), ‘A new objective method for scoring human sperm hyperactivation based on head axis angle deviation’, *International Journal of Andrology* **24**(4), 189–196.
- McDougall, S. R., Anderson, A. R. A. and Chaplain, M. A. J. (2006), ‘Mathematical modelling of dynamic adaptive tumour-induced angiogenesis: Clinical implications and therapeutic targeting strategies’, *Journal of Theoretical Biology* **241**(3), 564–589.
- McGaughey, R. W., Marston, J. H. and Chang, M. C. (1968), ‘Fertilizing life of mouse spermatozoa in the female tract’, *Journal of Reproductive Fertility* **16**(1), 147–150.
- Meyers, D., Skinner, S. and Sloan, K. (1992), ‘Surfaces from contours’, *ACM Transactions on Graphics* **11**(3), 228–258.
- Mi, Q., Rivière, B., Clermont, G., Steed, D. L. and Vodovotz, Y. (2007), ‘Agent-based model of inflammation and wound healing: insights into diabetic foot ulcer pathology and the role of transforming growth factor-1’, *Wound Repair and Regeneration* **15**(5), 671–682.
- Moore, H., Dvorakova, K., Jenkins, N. and Breed, W. (2002), ‘Exceptional sperm cooperation in the wood mouse’, *Nature* **418**(6894), 174–177.
- Moore, M. and Wilhelms, J. (1988), Collision detection and response for computer animationr3, in ‘SIGGRAPH ’88: Proceedings of the 15th annual conference on

- Computer graphics and interactive techniques', ACM, New York, NY, USA, pp. 289–298.
- Mortimer, S. T. (2000), 'CASA-practical aspects', *Journal of Andrology* **21**(4), 515–524.
- Murray, J. D. (2007), *Mathematical Biology: I. An Introduction (Interdisciplinary Applied Mathematics) (Pt. 1)*, Interdisciplinary applied mathematics, 3rd edn, Springer.
- Nedel, L. P. and Thalmann, D. (1998), Real time muscle deformations using Mass-Spring systems, in 'CGI '98: Proceedings of the Computer Graphics International 1998', IEEE Computer Society, Washington, DC, USA, pp. 156+.
- Nicol, A. and McLaren, A. (1974), 'An effect of the female genotype on sperm transport in mice', *Journal of Reproductive Fertility* **39**(2), 421–424.
- Nilsson, O. and Reinius, S. (1969), Light and electron microscopic structure of the oviduct, in E. S. E. Hafez and R. J. Blandau, eds, 'The Mammalian Oviduct', The University of Chicago Press, Chicago, chapter 3, pp. 57–84.
- Noble, D. (2002a), 'Modeling the heart—from genes to cells to the whole organ', *Science* **295**(5560), 1678–1682.
- Noble, D. (2002b), 'The rise of computational biology.', *Nature Reviews Molecular Cell Biology* **3**(6), 459–463.
- Noble, D. (2003), 'The future: putting Humpty-Dumpty together again', *Biochemical Society Transactions* **31**(1), 156–158.
- Noble, D. (2006), *The Music of Life: Biology Beyond Genes*, Oxford University Press, USA.
- NVIDIA, C. (2007), 'Compute Unified Device Architecture Programming Guide', *NVIDIA: Santa Clara, CA*.
- Nyland, L., Harris, M. and Prins, J. (2007), Fast N-Body simulation with CUDA, in H. Nguyen, ed., 'GPU Gems 3', Addison Wesley Professional, chapter 31.
- Olds, P. J. (1970), 'Effect of the t locus on sperm distribution in the house mouse', *Biology of Reproduction* **2**(1), 91–97.
- Oren-Benaroya, R., Kipnis, J. and Eisenbach, M. (2007), 'Phagocytosis of human post-capacitated spermatozoa by macrophages', *Human Reproduction* **22**(11), 2947–2955.

- Oshima, M., Torii, R., Kobayashi, T., Taniguchi, N. and Takagi, K. (2001), 'Finite element simulation of blood flow in the cerebral artery', *Computer Methods in Applied Mechanics and Engineering* **191**(6-7), 661–671.
- Perkins, J. L. (1972), Fluid flow of the oviduct, *in* A. D. Johnson and C. W. Foley, eds, 'The oviduct and its functions', Academic Press, New York, chapter 4, pp. 119–132.
- Phillips, D. M. (1972), 'Comparative analysis of mammalian sperm motility', *The Journal of Cell Biology* **53**(2), 561–573.
- Poddar, A. H., Krol, A., Beaumont, J., Price, R. L., Slamani, M. A., Fawcett, J., Subramanian, A., Coman, I. L., Lipson, E. D., Lipson, E. D. and Feiglin, D. H. (2005), Ultrahigh resolution 3D model of murine heart from micro-CT and serial confocal laser scanning microscopy images, *in* 'Nuclear Science Symposium and Medical Imaging Conference', Vol. 5, IEEE, pp. 2615–2617.
- Pop, M. and Salzberg, S. L. (2008), 'Bioinformatics challenges of new sequencing technology.', *Trends in genetics : TIG* **24**(3), 142–149.
- Qiang, W., Pan, Z., Chun, C. and Jianjun, B. (2007), 'Surface rendering for parallel slices of contours from medical imaging', *Computing in Science and Engineering* **9**(1), 32–37.
- Railsback, S. F., Lytinen, S. L. and Jackson, S. K. (2006), 'Agent-based simulation platforms: Review and development recommendations', *Simulation* **82**(9), 609–623.
- Reed, K. L., Rose, K. A. and Whitmore, R. C. (1984), 'Latin hypercube analysis of parameter sensitivity in a large model of outdoor recreation demand', *Ecological Modelling* **24**(3-4), 159–169.
- Reeves, W. T. (1983), 'Particle systems - a technique for modeling a class of fuzzy objects', *ACM Transactions on Graphics* **2**(2), 91–108.
- Richmond, P., Coakley, S. and Romano, D. M. (2009), A high performance agent based modelling framework on graphics card hardware with CUDA, *in* 'AAMAS '09: Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems', International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, pp. 1125–1126.
- Richmond, P. and Romano, D. (2011), Template-Driven Agent-Based modeling and simulation with CUDA, *in* W.-m. W. Hwu, ed., 'GPU Computing Gems Emerald Edition', 1 edn, Applications of GPU Computing Series, Morgan Kaufmann, chapter 21, pp. 313–324.

- Riffell, J. A. and Zimmer, R. K. (2007), ‘Sex and flow: the consequences of fluid shear for sperm egg interactions’, *Journal of Experimental Biology* **210**(20), 3644–3660.
- ROAD, R. (2005), ‘Recursive porous agent simulation toolkit (repast)’. [Last Accessed: 26-September-2011].  
**URL:** <http://repast.sourceforge.net>
- Rothschild (1963), ‘Non-random distribution of bull spermatozoa in a drop of sperm suspension’, *Nature* **198**(4886), 1221–1222.
- Rouwé, J. (2003), ‘Collision detection with swept spheres and ellipsoids’.  
**URL:** <http://www.jrouwe.nl/sweptellipsoid/SweptEllipsoid.pdf>
- Salekdeh, G. H. and Komatsu, S. (2007), ‘Crop proteomics: Aim at sustainable agriculture of tomorrow’, *Proteomics* **7**(16), 2976–2996.
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M. and Tarantola, S. (2008), *Global Sensitivity Analysis: The Primer*, 1 edn, Wiley-Interscience.
- Samsonova, A. A., Niranjan, M., Russell, S. and Brazma, A. (2007), ‘Prediction of gene expression in embryonic structures of drosophila melanogaster’, *PLoS Computational Biology* **3**(7), e144+.
- Sanders, J. and Kandrot, E. (2010), *CUDA by Example: An Introduction to General-Purpose GPU Programming*, 1 edn, Addison-Wesley Professional.
- Sargent, R. G. (2005), Verification and validation of simulation models, in ‘Proceedings of the 37th conference on Winter simulation’, WSC ’05, Winter Simulation Conference, pp. 130–143.
- Sargent, R. G. (2007), Verification and validation of simulation models, in ‘WSC ’07: Proceedings of the 39th conference on Winter simulation’, IEEE Press, Piscataway, NJ, USA, pp. 124–137.
- Sauer, U., Heinemann, M. and Zamboni, N. (2007), ‘Getting closer to the whole picture’, *Science* **316**(5824), 550–551.
- Sauro, H. M., Harel, D., Kwiatkowska, M., Shaffer, C. A., Uhrmacher, A. M., Hucka, M., Mendes, P., Strömback, L. and Tyson, J. J. (2006), Challenges for modeling and simulation methods in systems biology, in ‘WSC ’06: Proceedings of the 38th conference on Winter simulation’, Winter Simulation Conference, pp. 1720–1730.

- Scott, A. (2004), 'Reductionism revisited', *Journal of Consciousness Studies* **11**(2), 51–68.
- Segal, E., Raveh-Sadka, T., Schroeder, M., Unnerstall, U. and Gaul, U. (2008), 'Predicting expression patterns from regulatory sequence in drosophila segmentation', *Nature* **451**(7178), 535–540.
- Segovia-Juarez, J. L., Ganguli, S. and Kirschner, D. (2004), 'Identifying control mechanisms of granuloma formation during m. tuberculosis infection using an agent-based model', *Journal of Theoretical Biology* **231**(3), 357–376.
- Seo, D.-b., Agca, Y., Feng, Z. and Critser, J. (2007), 'Development of sorting, aligning, and orienting motile sperm using microfluidic device operated by hydrostatic pressure', *Microfluidics and Nanofluidics* **3**(5), 561–570.
- Seytanoglu, A., Georgiou, A. S., Sostaric, E., Watson, P. F., Holt, W. V. and Fazeli, A. (2008), 'Oviductal cell proteome alterations during the reproductive cycle in pigs', *Journal of Proteome Research* **7**(7), 2825–2833.
- Sharpe, R. M. (2000), 'Lifestyle and environmental contribution to male infertility', *British Medical Bulletin* **56**(3), 630–642.
- Shendarkar, A., Vasudevan, K., Lee, S. and Son, Y.-J. (2006), Crowd simulation for emergency response using BDI agent based on virtual reality, in 'Simulation Conference, 2006. WSC 06. Proceedings of the Winter', IEEE, pp. 545–553.
- Shreedhara, K. S. and Indira, S. P. (2006), Construction of 3-D objects using 2-D cross sectional data and curves, in '2006 International Conference on Advanced Computing and Communications', IEEE, pp. 630–631.
- Siddiquey, A. K. S. and Cohen, J. (1982), 'In-vitro fertilization in the mouse and the relevance of different sperm/egg concentrations and volumes', *Journal of Reproduction and Fertility* **66**(1), 237–242.
- Slager, C. J., Wentzel, J. J., Schuurbiers, J. C. H., Oomen, J. A. F., Kloet, J., Krams, R., von Birgelen, C., van der Giessen, W. J., Serruys, P. W. and de Feyter, P. J. (2000), 'True 3-Dimensional reconstruction of coronary arteries in patients by fusion of angiography and IVUS (ANGUS) and its quantitative validation', *Circulation* **102**(5), 511–516.
- Smirnov, A. (2003), 'Tool-assisted mesh generation based on a tissue-growth model', *Medical and Biological Engineering and Computing* **41**(4), 494–497.



- Smith, A. E., Slepchenko, B. M., Schaff, J. C., Loew, L. M. and Macara, I. G. (2002), ‘Systems analysis of ran transport.’, *Science (New York, N.Y.)* **295**(5554), 488–491.
- Smith, A. L. and Lodge, J. R. (1987), ‘Interactions of aged gametes: In vitro fertilization using in vitro-aged sperm and in vivo-aged ova in the mouse’, *Gamete Research* **16**(1), 47–56.
- Smith, D. J., Gaffney, E. A., Blake, J. R. and Kirkman-brown, J. C. (2009), ‘Human sperm accumulation near surfaces: a simulation study’, *Journal of Fluid Mechanics* **621**(4), 289–320.
- Smith, T. T., Koyanagi, F. and Yanagimachi, R. (1987), ‘Distribution and number of spermatozoa in the oviduct of the golden hamster after natural mating and artificial insemination.’, *Biology of Reproduction* **37**(1), 225–234.
- Smith, T. T. and Yanagimachi, R. (1990), ‘The viability of hamster spermatozoa stored in the isthmus of the oviduct: the importance of sperm-epithelium contact for sperm survival.’, *Biology of reproduction* **42**(3), 450–457.
- Smith, T. T. and Yanagimachi, R. (1991), ‘Attachment and release of spermatozoa from the caudal isthmus of the hamster oviduct’, *Journal of Reproduction and Fertility* **91**(2), 567–573.
- Ŝolín, P. (2005), *Partial Differential Equations and the Finite Element Method (Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts)*, 1 edn, Wiley-Interscience.
- Sostaric, E., Georgiou, A. S., Wong, C. H., Watson, P. F., Holt, W. V. and Fazeli, A. (2006), ‘Global profiling of surface plasma membrane proteome of oviductal epithelial cells’, *Journal of Proteome Research* **5**(11), 3029–3037.
- Southern, J., Pitt-Francis, J., Whiteley, J., Stokeley, D., Kobashi, H., Nobes, R., Kadooka, Y. and Gavaghan, D. (2008), ‘Multi-scale computational modelling in biology and physiology’, *Progress in Biophysics and Molecular Biology* **96**(1-3), 60–89.
- Stahl, D., Ezquerra, N. and Turk, G. (2002), Bag-of-particles as a deformable model, in ‘VISSYM ’02: Proceedings of the symposium on Data Visualisation 2002’, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, pp. 141–150.
- Steve Railsback, S. L. and Grimm, V. (2005), ‘Stupidmodel and extensions: A template and teaching tool for agent-based modeling platforms’. [Last Accessed: 26-September-2011].
- URL:** <http://condor.depaul.edu/šlytinen/abm/StupidModelFormulation.pdf>

- Stockard, C. R. and Papanicolaou, G. N. (1917), ‘The existence of a typical oestrous cycle in the guinea-pig-with a study of its histological and physiological changes’, *American Journal of Anatomy* **22**(2), 225–283.
- Suarez, S. S. (1987), ‘Sperm transport and motility in the mouse oviduct: observations in situ.’, *Biology of reproduction* **36**(1), 203–210.
- Suarez, S. S. (2002), ‘Formation of a reservoir of sperm in the oviduct’, *Reproduction in Domestic Animals* **37**(3), 140–143.
- Suarez, S. S. (2008a), ‘Control of hyperactivation in sperm’, *Human Reproduction Update* **14**(6), 647–657.
- Suarez, S. S. (2008b), ‘Regulation of sperm storage and movement in the mammalian oviduct.’, *The International Journal of Developmental Biology* **52**(5-6), 455–462.
- Suarez, S. S. and Dai, X. (1992), ‘Hyperactivation enhances mouse sperm capacity for penetrating viscoelastic media.’, *Biology of Reproduction* **46**(4), 686–691.
- Suarez, S. S., Katz, D. F., Owen, D. H., Andrew, J. B. and Powell, R. L. (1991), ‘Evidence for the function of hyperactivated motility in sperm.’, *Biology of Reproduction* **44**(2), 375–381.
- Suarez, S. S. and Pacey, A. A. (2006), ‘Sperm transport in the female reproductive tract’, *Human Reproduction Update* **12**(1), 23–37.
- Sun, F., Bahat, A., Gakamsky, A., Girsh, E., Katz, N., Giojalas, L. C., Tur-Kaspa, I. and Eisenbach, M. (2005), ‘Human sperm chemotaxis: both the oocyte and its surrounding cumulus cells secrete sperm chemoattractants’, *Human Reproduction* **20**(3), 761–767.
- Swan, S. H., Elkin, E. P. and Fenster, L. (1997), ‘Have sperm densities declined? a reanalysis of global trend data.’, *Environmental health perspectives* **105**(11), 1228–1232.
- SWARM (2009), ‘Swarm’. [Last Accessed: 26-September-2011].  
**URL:** <http://www.swarm.org>
- Szeliski, R. and Tonnesen, D. (1992), ‘Surface modeling with oriented particle systems’, *ACM SIGGRAPH Computer Graphics* **26**(2), 185–194.
- Taylor, G. (1951), ‘Analysis of the swimming of microscopic organisms’, *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **209**(1099), 447–461.

- Tessler, S. and Olds-Clarke, P. (1981), ‘Male genotype influences sperm transport in female mice’, *Biology of Reproduction* **24**(4), 806–813.
- Tessler, S. and Olds-Clarke, P. (1985), ‘Linear and nonlinear mouse sperm motility patterns. a quantitative classification’, *Journal of Andrology* **6**(1), 35–44.
- Thibault, W. C. and Naylor, B. F. (1987), ‘Set operations on polyhedra using binary space partitioning trees’, *ACM SIGGRAPH Computer Graphics* **21**(4), 153–162.
- Thorne, B. C., Bailey, A. M. and Peirce, S. M. (2007), ‘Combining experiments with multi-cell agent-based modeling to study biological tissue patterning’, *Briefings in Bioinformatics* **8**(4), 245–257.
- Tienthai, P., Johannisson, A. and Rodriguez-Martinez, H. (2004), ‘Sperm capacitation in the porcine oviduct’, *Animal Reproduction Science* **80**(1-2), 131–146.
- Tillier, Y., Paccini, A., Durand-Reville, M., Bay, F. and Chenot, J. L. (2003), ‘Three-dimensional finite element modelling for soft tissues surgery’, *International Congress Series* **1256**, 349–355.
- Topping, C. J., Hansen, T. S., Jensen, T. S., Jepsen, J. U., Nikolajsen, F. and Odderskær, P. (2003), ‘ALMaSS, an agent-based model for animals in temperate european landscapes’, *Ecological Modelling* **167**(1-2), 65–82.
- Treece, G. M., Prager, R. W., Gee, A. H. and Berman, L. (2000), ‘Surface interpolation from sparse cross sections using region correspondence.’, *IEEE transactions on medical imaging* **19**(11), 1106–1114.
- Tsunoda, Y. and Chang, M. C. (1975), ‘Penetration of mouse eggs in vitro: Optimal sperm concentration and minimal number of spermatozoa’, *Journal of Reproduction and Fertility* **44**(1), 139–142.
- Van Soom, A., Vandaele, L., Peelman, L. J., Goossens, K. and Fazeli, A. (2010), ‘Modeling the interaction of gametes and embryos with the maternal genital tract: From in vivo to in silico’, *Theriogenology* **73**(6), 828–837.
- Vignon-Clementel, I. E., Alberto Figueroa, C., Jansen, K. E. and Taylor, C. A. (2006), ‘Outflow boundary conditions for three-dimensional finite element modeling of blood flow and pressure in arteries’, *Computer Methods in Applied Mechanics and Engineering* **195**(29-32), 3776–3796.
- Walker, D. C., Hill, G., Wood, S. M., Smallwood, R. H. and Southgate, J. (2004), ‘Agent-Based computational modeling of wounded epithelial cell monolayers’, *Transactions on Nanobioscience* **3**(3), 153–163.

- Walker, D. C. and Southgate, J. (2009), ‘The virtual cella candidate co-ordinator for ‘middle-out’ modelling of biological systems’, *Briefings in Bioinformatics* **10**(4), 450–461.
- Wang, Z., Birch, C. M., Sagotsky, J. and Deisboeck, T. S. (2009), ‘Cross-scale, cross-pathway evaluation using an agent-based non-small cell lung cancer model’, *Bioinformatics* **25**(18), 2389–2396.
- Wang, Z., Zhang, L., Sagotsky, J. and Deisboeck, T. S. (2007), ‘Simulating non-small cell lung cancer with a multiscale agent-based model.’, *Theoretical biology & medical modelling* **4**(1), 50+.
- Wanggren, K., Stavreus-Evers, A., Olsson, C., Andersson, E. and Gemzell-Danielsson, K. (2008), ‘Regulation of muscular contractions in the human fallopian tube through prostaglandins and progestagens’, *Human Reproduction* **23**(10), 2359–2368.
- Watt, A. H. (1999), *3D Computer Graphics (3rd Edition)*, Addison Wesley.
- Watt, A. and Policarpo, F. (2000), *3D Games: Real-Time Rendering and Software Technology, Volume 1 (With CD-ROM)*, Addison Wesley.
- Werner, R., Ehrhardt, J., Schmidt, R. and Handels, H. (2009), ‘Patient-specific finite element modeling of respiratory lung motion using 4D CT image data’, *Medical Physics* **36**(5), 1500–1511.
- Wilensky, U. (1999), ‘Netlogo’. [Last Accessed: 26-September-2011].  
**URL:** <http://ccl.northwestern.edu/netlogo/>
- Wilke, A. (2003), ‘Bioinformatics support for high-throughput proteomics’, *Journal of Biotechnology* **106**(2-3), 147–156.
- Williams, M., Hill, C. J., Scudamore, I., Dunphy, B., Cooke, I. D. and Barratt, C. L. R. (1993), ‘Physiology: Sperm numbers and distribution within the human fallopian tube around ovulation’, *Human Reproduction* **8**(12), 2019–2026.
- Wissel, C. (1992), ‘Aims and limits of ecological modelling exemplified by island theory’, *Ecological Modelling* **63**(1-4), 1–12.
- Wolf, D. (1978), ‘The block to sperm penetration in zona-free mouse eggs’, *Developmental Biology* **64**(1), 1–10.
- Wolfram, S. (1984), ‘Cellular automata as models of complexity’, *Nature* **311**, 419–424.
- Woolley, D. M. (2003), ‘Motility of spermatozoa at surfaces’, *Reproduction* **126**(2), 259–270.

- Yang, X., Dillon, R. and Fauci, L. (2008), ‘An integrative computational model of multiciliary beating’, *Bulletin of Mathematical Biology* **70**(4), 1192–1215.
- Yaniz, J. L., Lopez-Gatius, F. and Hunter, R. H. F. (2006), ‘Scanning electron microscopic study of the functional anatomy of the porcine oviductal mucosa’, *Anatomia, Histologia, Embryologia: Journal of Veterinary Medicine Series C* **35**(1), 28–34.
- Younglai, E. V., Holloway, A. C. and Foster, W. G. (2005), ‘Environmental and occupational factors affecting fertility and IVF success’, *Human Reproduction Update* **11**(1), 43–57.
- Zamboni, L. (1972), Fertilization in the mouse, *in* Kamran and E. S. E. Hafez, eds, ‘Biology of Mammalian Fertilization and Implantation’, Charles C Thomas, Springfield, Illinois, USA, chapter 8, pp. 213–262.
- Zana, J., Muffat-Joly, M., Thomas, D., Orfila, J., Salat-Baroux, J. and Pocidalo, J. J. (1991), ‘Roxithromycin treatment of mouse chlamydial salpingitis and protective effect on fertility.’, *Antimicrobial agents and chemotherapy* **35**(3), 430–435.
- Zervomanolakis, I., Ott, H. W., Hadziomerovic, D., Mattle, V., Seeber, B. E., Virgolini, I., Heute, D., Kissler, S., Leyendecker, G. and Wildt, L. (2007), ‘Physiology of upward transport in the human female genital tract’, *Annals of the New York Academy of Sciences* **1101**(Reproductive Biomechanics), 1–20.
- Zinzen, R. P., Senger, K., Levine, M. and Papatsenko, D. (2006), ‘Computational models for neurogenic gene expression in the drosophila embryo.’, *Current biology : CB* **16**(13), 1358–1365.
- Zwillinger, D. (1997), *Handbook of Differential Equations, Third Edition*, 3 edn, Academic Press.

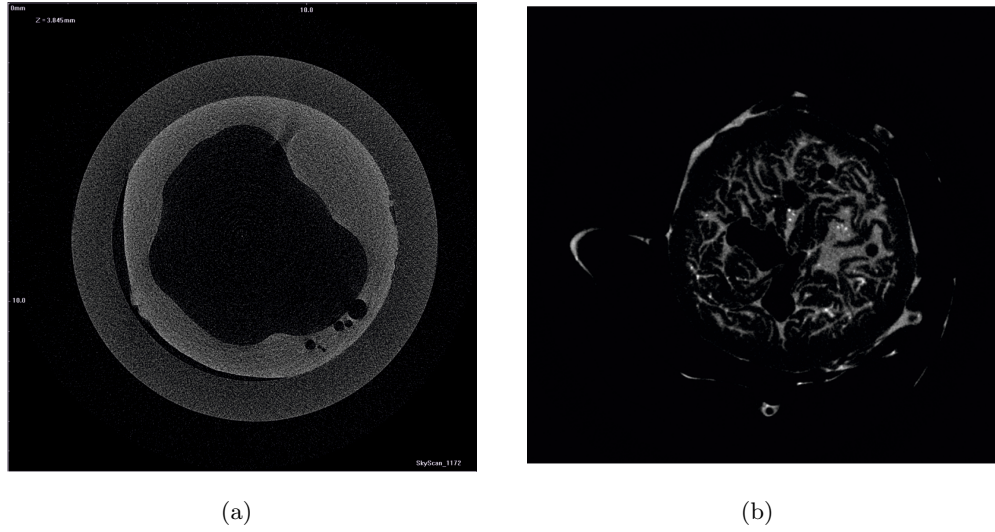
## Appendix A

### Micro-CT experiments

The experiments described in this section were carried out during my MSc project, reported in my MSc Thesis (Burkitt, 2008), and only included here for convenience.

Several tests were carried out using a SkyScan 1172 High Resolution Micro-CT scanner to see if any internal structure could be identified. A Micro-CT scanner works by projecting x-rays through a rotating object onto a sensor. Some of the rays pass through the object, and others are blocked. A set of images recording the absorbance is generated, and from this, cross sectional images are created. To use Micro-CT imaging for development of 3D models of oviduct, porcine oviducts, both fresh and fixed in formaldehyde were cut into  $3\text{cm}$  segments. For the first set of tests using the fresh oviduct, scans were performed on the oviduct unmodified, then filled with air, distilled water and then a calcium chloride solution using different resolution settings. For all tests, no internal structure was visible. Figure A.1(a) shows an example of an oviduct filled with air captured using  $100\text{Kv}$  and  $100\mu\text{A}$  at a resolution of  $7.86\mu\text{m}$  per pixel. The second set of tests using the fixed oviduct was carried out using a strong contrast agent. The oviduct was washed with Phosphate Buffer Saline (PBS) and then injected with a Barium Sulphate solution, mixed at a concentration of 70%. The results from this scan were a great improvement on any of the previous scans.

Figure A.1(b) shows an example of an oviduct filled with Barium Sulphate Solution captured using  $49\text{kV}$  and  $200\mu\text{A}$  at a resolution of  $9.96\mu\text{m}$  per pixel. Unlike the previous images, the tissue of the oviduct looked dark, and the light areas were where the Barium Sulphate had filled the structure. The Barium Sulphate blocks more of the radiation, and therefore looked much lighter in comparison to the tissue. Some of the fine detail of the internal structure was clearly visible, although when compared to slides obtained from the histology; these images looked quite different and were not very detailed.



*Figure A.1:* A porcine oviduct scanned using a Micro-CT scanner. (a) the oviduct filled with air scanned without contrasting agent and (b) the oviduct segment scanned with Barium Sulphate solution contrasting agent.

However, several limitations of this approach were identified. The Micro-CT scanner was able to process small sections at a time, so multiple scans on sequential sections needed to be performed. In addition the act of cutting these sections caused the internal structure to be damaged and distorted. As only the Barium Sulphate solution is being detected, and the tissue is identified as an absence of Barium Sulphate, it is highly likely that internal structure captured will not be complete, as there is no guarantee that the internal surface of all folds will be completely covered. Finally, the act of injecting the solution into the oviduct disrupts the internal structure, causing artefacts which would not appear in the real organ. As several sections will need to be combined together to construct a complete 3D model, it is likely that this disruption will cause non-continuity between sections, making it difficult to construct a continuous and complete 3D model.

## Appendix B

### Obtaining histology images from mouse oviducts

Due to the regulations surrounding animal experimentation and the training required, my role in the experiments described in this section was to observe, assist with the histology staining and record measurements only, not to handle the mice.

In order to construct a more realistic internal and external structure for the oviduct, a new set of histology images was required. The oestrous cycle in mice is between four and five days long (Johnson and Everitt, 1999). However, determining the current stage of that cycle an individual mouse is a non-trivial process. An experiment was set up to first identify the mice in either the oestrous or diestrus phase, and then harvest the oviducts from the mice to obtain an equal number of oestrous and diestrus oviducts.

Sixty female mice were used for the experiment. In the absence of male mice, females remain fixed in the luteal phase of the oestrous cycle (Johnson and Everitt, 1999). To induce ovulation without the possibility of the mice becoming pregnant, droppings from male mice were planted in the bedding of the females. Before the identification process began, the mice were left for one week to acclimatise, with the male droppings introduced.

Two different methods were initially investigated to determine the phase of the oestrous cycle. The first method, initially proposed by Champlin et al. (1973), attempts to determine the phase of the oestrous cycle in mice by the appearance of the vaginal alone. The second method, originally described by Stockard and Papanicolaou (1917) in guinea-pigs, requires a swab to be taken from the mouse vagina, smeared onto a slide, stained and then looked at under a microscope. A large number of cells indicated the oestrous cycle, and very few cells indicates the diestrus cycle.

Initially, both methods were used to see if they produced consistent results. However,



the first method was abandoned, as the changes in the appearance of the vagina were too subtle to consistently distinguish with any certainty, and the swab and stain method was followed for the remainder of the experiment.

The experiment was performed over five days. Mice that were either definitely oestrous or definitely diestrus were harvested, and their oviducts stored. Mice where the phase could not be determined with sufficient certainty were left and checked the following day. As oestrous is more difficult to detect than diestrus, when the number of diestrus mice was greater than the number of oestrous mice, the surplus diestrus mice were returned, thus ensuring a consistent number of oestrous and diestrus oviducts. On day five, all but seven of the mice could be identified as either oestrous or diestrus, so all these mice were harvested, the remaining seven mice were discarded and the experiment was terminated. The end result was forty two oestrous oviducts and sixty two diestrus oviducts.

## Appendix C

### Oviduct model details

#### Oviduct measurements

The table below shows the final measurement values used to generate the 3D environments. The Region column shows which section of the tube each set of measurements relates to. The Position Fraction column shows the position along the oviductal tube as a fraction, with 0 at the isthmus and 1 at the end of the ampulla.

Region	Position Fraction	Inner Radius Long ( $\mu m$ )	Inner Radius Short ( $\mu m$ )	Outer Radius Long ( $\mu m$ )	Lumen Distance Long ( $\mu m$ )	Lumen Distance Short ( $\mu m$ )	Inner Length ( $\mu m$ )
<b>Sperm Reservoir</b>							
Intramural	0.0000	157.9778	119.0557	269.3472	17.4004	18.6204	1515.0165
	0.0154	171.6180	124.4101	278.3054	17.9940	19.6042	1649.3843
	0.0308	184.4588	130.9395	279.0686	20.5962	21.2321	1779.0033
	0.0462	189.9537	136.0637	276.7573	24.8792	22.4436	1838.1823
	0.0615	181.7647	135.9062	266.8360	25.0997	23.7570	1777.0796
	0.0769	179.9536	142.2054	264.7100	24.5406	24.8940	1767.7882
Extramural	0.0923	175.9851	149.0086	261.3084	25.4647	26.0789	1725.2683
	0.1077	167.7646	150.6149	251.0091	29.6222	26.6458	1585.8031
	0.1231	157.2527	145.0637	237.1901	30.6852	25.9371	1405.5660
	0.1385	135.9928	122.9534	208.5601	28.9844	25.3702	1131.7379
	0.1538	117.9927	101.9769	184.1348	26.7167	25.7482	925.3746
<b>Lower Isthmus</b>							
	0.1692	99.2131	78.2366	166.7725	24.9450	27.8742	730.4918
	0.1846	94.1107	75.6854	174.4261	27.2893	29.7394	680.6018
	0.2000	92.7946	78.5201	181.8839	30.0045	30.2725	678.1721
	0.2154	93.1883	82.1421	181.8052	33.5368	29.6615	675.8099
	0.2308	92.6214	77.9531	174.6240	32.1201	28.3380	652.6916
	0.2462	91.1028	75.8271	173.3079	29.4049	26.1797	676.3813

Region	Position Fraction	Inner Radius Long ( $\mu m$ )	Inner Radius Short ( $\mu m$ )	Outer Radius Long ( $\mu m$ )	Lumen Distance Long ( $\mu m$ )	Lumen Distance Short ( $\mu m$ )	Inner Length ( $\mu m$ )
	0.2615	87.1658	73.6224	165.5914	26.3686	22.3403	648.9796
	0.2769	79.3705	68.0318	158.0323	22.8722	23.2300	580.3966
	0.2923	72.8153	59.8822	141.3786	21.9509	24.6509	465.5928
	0.3077	68.5633	55.3467	132.8747	22.5887	27.1171	426.4745
	0.3231	71.7017	51.3985	122.9534	25.3196	26.0905	434.6748
	0.3385	75.2197	50.9429	124.7250	26.9667	26.4413	448.4432
	0.3538	76.0465	51.3445	133.5833	25.8092	26.4224	455.6715
	0.3692	72.4357	53.8754	141.8511	22.3472	23.7167	456.6839
<b>Upper Isthmus</b>							
Upper Isthmus	0.3846	78.3075	58.2286	155.3157	25.5190	23.3340	507.4041
	0.4000	86.8351	59.8113	163.4654	29.7946	23.3529	559.7035
	0.4154	97.2288	60.5200	170.7882	34.3702	25.6537	606.4754
	0.4308	97.0871	53.2916	164.9772	28.3353	23.4256	583.3729
	0.4462	96.0342	48.7561	160.1178	23.4475	22.7372	567.4179
	0.4615	85.1680	46.1577	148.7791	19.1128	21.5844	508.3625
	0.4769	76.2546	47.3231	141.9130	19.0501	21.3369	470.7560
	0.4923	71.0240	48.0317	139.2133	18.5544	20.2536	447.4038
	0.5077	79.7642	47.7955	147.0086	17.7371	17.9056	506.4592
	0.5231	82.7248	49.7483	151.0401	18.8883	19.0432	531.5932
	0.5385	80.0083	51.5199	150.0952	20.3741	21.3818	520.7270
	0.5538	90.4020	63.8035	160.7252	20.0410	23.2385	687.9719
	0.6000	109.2525	70.0397	171.0717	18.5812	21.1664	870.2405
Ampullar-Isthmic Junction	0.6154	132.1829	74.4436	182.6128	17.5000	18.0118	1062.3899
	0.6308	135.0175	83.8924	180.2506	20.8590	16.8213	1516.2257
	0.6615	140.9231	97.1208	187.9987	24.1756	17.5621	1763.1245
	0.6769	157.0874	114.0950	191.1269	25.5592	17.9339	2084.0418
<b>Ampulla</b>							
Ampulla Start	0.6923	180.2371	118.3470	219.6153	24.6644	16.2898	2099.5379
	0.7077	184.2529	115.7486	242.3634	25.0867	18.2505	2157.6484
	0.7385	177.8749	110.3155	226.3003	24.2987	20.0883	2031.5061
	0.7538	168.4260	101.3391	205.5128	28.7718	21.7324	1783.0009
Site of Fertilisation	0.7692	266.8596	193.3238	314.6118	46.9845	30.0946	3050.3063
	0.8000	303.2802	224.6184	353.1206	56.3431	29.5394	3880.5498
	0.8154	338.8623	270.9439	393.2984	59.0346	29.7189	4938.2818
	0.8308	302.7538	245.3681	356.0292	35.4843	23.2696	4851.5888
	0.8462	297.0420	248.0894	346.1819	32.6907	23.8248	5115.3980
Ampulla End	0.8615	237.2813	164.6937	298.6917	31.7851	23.1216	5141.6622
	0.8769	253.6312	180.2844	287.2298	32.0946	23.0080	5253.8511
	0.8923	240.8752	173.1977	280.7730	26.4757	19.1340	4942.3850
	0.9077	220.7491	164.4103	265.0407	20.5938	19.7151	4497.1166

Region	Position Fraction	Inner Radius Long ( $\mu m$ )	Inner Radius Short ( $\mu m$ )	Outer Radius Long ( $\mu m$ )	Lumen Distance Long ( $\mu m$ )	Lumen Distance Short ( $\mu m$ )	Inner Length ( $\mu m$ )
	0.9231	193.4655	157.3236	240.2374	15.0946	16.5828	3984.9645
	0.9385	157.4181	128.8825	194.8356	18.1938	18.1418	3643.4825
	0.9538	160.2528	127.7487	198.8041	20.6033	17.3812	3581.1200
<b>Pre-Ampulla</b>							
	0.9692	159.6858	118.3943	200.7884	22.5592	18.4725	3279.2287
	0.9846	152.8826	111.9691	194.3632	22.0867	17.9528	3031.6685
	1.0000	147.6858	106.7249	189.4970	21.9686	19.4883	2829.6990

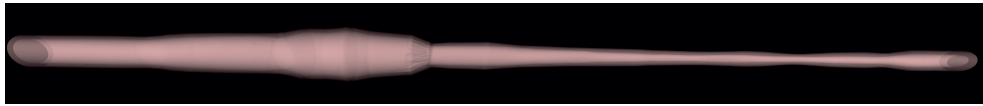
### Oviduct model variants

The section provides details of the final model variants used within the computational simulations. Figures C.1 and C.2 show how the different curve forms appear with and without internal folds, and with or without bends and internal folds.

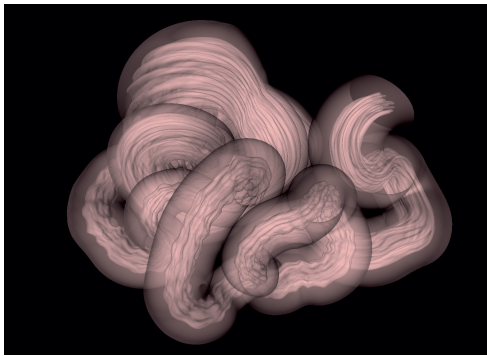
Env.	Fold	Curve	ID	Env.	Fold	Curve	ID
Normal	Oviduct A	Curve 1	OvA1	Scaled 90%	Oviduct A	Curve 1	OvA1_90
	Oviduct A	Curve 2	OvA2	Scaled 80%	Oviduct A	Curve 1	OvA1_80
	Oviduct A	Curve 3	OvA3	Modified	No Folds	Curve 1	OvNF1
	Oviduct B	Curve 1	OvB1		No Folds	Curve 2	OvNF2
	Oviduct B	Curve 2	OvB2		No Folds	Curve 3	OvNF3
	Oviduct B	Curve 3	OvB3		Oviduct A	Straight	OvAS
	Oviduct C	Curve 1	OvC1		Oviduct B	Straight	OvBS
	Oviduct C	Curve 2	OvC2		Oviduct C	Straight	OvCS
	Oviduct C	Curve 3	OvC3		No Folds	Straight	OvNFS



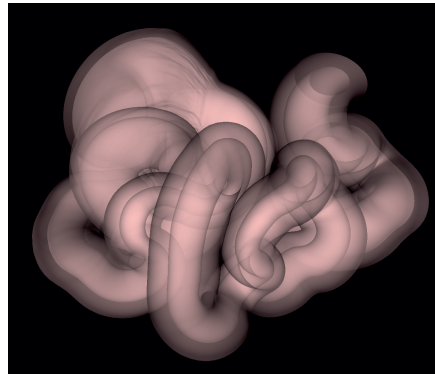
(a)



(b)

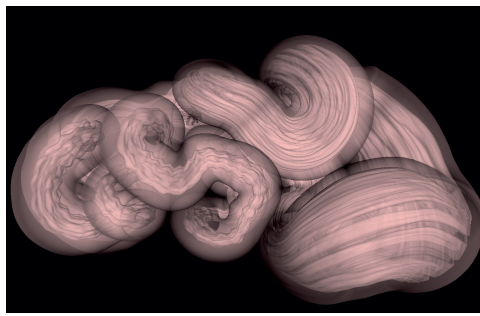


(c)

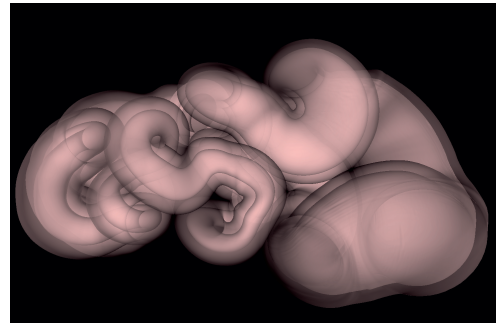


(d)

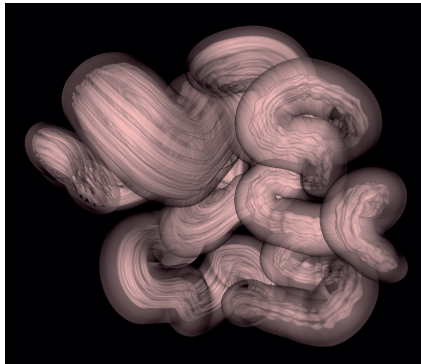
Figure C.1: (a) OvAS (b) OvNFS (c) OvA1 (d) OvNF1



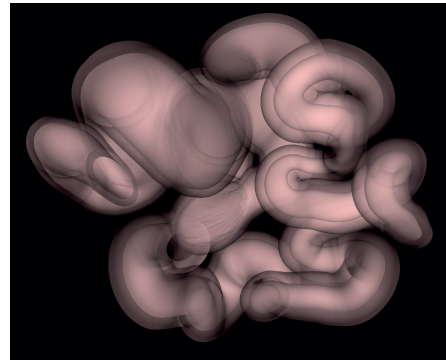
(a)



(b)



(c)



(d)

Figure C.2: (a) OvA2 (b) OvNF2 (c) OvA3 (d) OvNF3

## Appendix D

### Model assumptions

This section lists all the assumptions made during the development of the model. The assumptions prefixed with a G relate to the generic behaviour model, the assumptions prefixed with an E relate to the 3D environment and assumptions prefixed with a M relate to the mouse specific behaviour model.

ID	Assumption
G1	Capacitation can be represented as a probability and is not related to the number of sperm in the oviduct
G2	Sperm are unable to detach from the oviduct epithelium until they become capacitated
G3	All sperm agents are initially uncapacitated
G4	All sperm live for the same amount of time once capacitated, and amount of time attached to the epithelium, amount of time swimming non-progressively do not alter this.
G5	All sperm in the oviduct have the same motility and are capable of fertilisation
G6	All progressive swimming sperm move forward
G7	All non-progressive swimming sperm move in the same manner
G8	Switching between progressive and non-progressive swimming is random and not based on any biological interactions
G9	All sperm swim at the same speed
M1	Only non-progressive sperm can detach from the oviduct
M2	Non-progressive sperm cannot attach to the oviduct
M3	Progressive sperm cannot detach from the oviduct
G10	Sperm which become capacitated later in the simulation will have the same lifetime and swimming characteristics as sperm which became capacitated at the beginning of the simulation
G11	Before oocytes have undergone the maturation process, sperm will not interact with them
G12	Sperm attachment to oocytes is not altered after the block to polyspermy is established
G13	Interactions between individual sperm in the oviduct are not likely or significant
M4	The narrow lumen makes it difficult for sperm to turn around
M5	Sperm will not start swimming until 1 hours <i>post coitus</i> , after the UTJ has closed

ID	Assumption
G14	Sperm and oocytes all have a similar size
G15	The fluid within the oviduct is static, or the flow is too weak to significantly influence the movement of sperm
E1	Recreating the complexity of the internal folds is sufficient to simulate the interactions that occur between sperm and the environment in vivo
E2	Three replicates of the external bends of the oviduct is sufficient to generalise for all oviducts
G16	Sperm swim at the same speed regardless of the current oviduct region
G17	Sperm have the same probability of attachment regardless of the current oviduct region
G18	Sperm penetrate the Zona Pellucida very quickly
G19	All sperm are initially distributed along the walls of the oviduct in the sperm reservoir
G20	All oocytes have a fixed position at the site of fertilisation in the ampulla and do not move until after the simulation finishes
G21	Sperm which reach the end of the pre-ampulla will leave and be unable to re-enter
E3	The structure of the oviduct will not change during the simulation
G22	Sperm interactions with the oviduct are limited to attachment and detachment only
G23	Sperm are not attracted to temperature changes or chemicals within the oviductal fluid
E4	The fine surface features of the oviduct, such as the crypts and cilia do not have a significant impact on the behaviour of sperm
G24	Sperm are not able to swim back out of the oviduct into the uterus
G25	The sperm population in the oviduct is static, and sperm in the uterus are not able to enter the oviduct during the simulation